

RESULT	15			
014898				
ID	014898	PRELIMINARY;	PRT;	506 AA.
OC	014898:			
DT	01-JAN-1998	(TEMBLrel. 05, Created)		
DT	01-JAN-1998	(TEMBLrel. 05, Last sequence update)		
DT	01-MAY-1999	(TEMBLrel. 10, Last annotation update)		
DE	ZINC FINGER PROTEIN (FRAGMENT).			
GN	HZF6.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;			
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
UN	[1]			
..	SEQUENCE FROM N.A.			
TX	MEDLINE; 98234550.			
RA	SHANNON M., STUBBS L.;			
RT	"Analysis of homologous XRCC1-linked zinc-finger gene families in			
RT	human and mouse: evidence for orthologous genes.";			
RL	Genomics 49:112-121(1998).			
DR	EMBL; AF027513; AAD12728.1; -.			
DR	PFAM; PF01952; KRAB; 1.			
DR	PFAM; PF00096; zf-C2H2; 8.			
DR	PROSITE; PS00028; ZINC_FINGER_C2H2; 6.			
KW	zinc-finger; Metal-binding; DNA-binding.			
FT	NON_TER 506			
SQ	SEQUENCE 506 AA; 57498 MW; 7A530D9B CRC32;			


```

[1] RN SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA BEUTLER E., GELBART T., WEST C., KUHL W., LEE P.L.;
RL Blood Cells Mol. Dis. 21:206-216(1995).
[2] RN
RC SEQUENCE FROM N.A.
RP TISSUE=OVARY;
RA LEE P.L.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
[3] RN
RC SEQUENCE FROM N.A.
RX MEDLINE: 97386587.
RA LEE P.L., GELBART T., WEST C., ADAMS M., BLACKSTONE R., BEUTLER E.;
RL "Three genes encoding zinc finger proteins on human chromosome
6p21.3: members of a new subclass of the Kruppel gene family
containing the conserved SCAN box domain.";
RL Genomics 43:191-201(1997).
DR EMBL: U57796; AAC02260.1; -.
DR EMBL: U88080; AAC51656.1; -.
DR EMBL: U88079; AAC51656.1; JOINED.
DR PFAM: PF00096; zf-C2H2; 9.
KW PROSITE: PS00028; ZINC_FINGER_C2H2; 9.
KR Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 578 AA; 65757 MW; COA9AD2B CRC32;

Query Match      25.9%; Score 291.5; DB 4; Length 578;
Best Local Similarity 40.3%; Pred. No. 7e-21;
Matches 56; Conservative 24; Mismatches 54; Indels 5; Gaps

QY 19 EIPIYSYSREYN-ENIKLERHYVSDSRPTSGKMNCDCVLGSCISFNVLWVKRSHTG 77
Db 431 ERPY-----ECNECGKAFSSHLLIGHORIHTGEKPYECDCEGKTFRSSHLLIGHORSHTG 486

QY 78 ERPFCNOCGASFTQKNLRLHKLTGTGPKFKCHLCNVACORRDALTGHRLTHSVSEKPY 137
Db 487 ERPKYCNCGRFAFSOKSLIEHORIHITGERPKYCKEKGAFNGMTGLIQHLRIHTGEKPY 546

QY 138 KCEFGCRSYKORSSLEEHK 156
Db 547 QCNECGKAFTQRSSLIRHQ 565

RESULT 10
P70590 PRELIMINARY; PRT; 591 AA.
ID P70590
AC P70590;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DI 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE KRAB-ZINC FINGER PROTEIN KZF-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1] RN
RC SEQUENCE FROM N.A.
RA BELLEFROID E.J., BOURGUIGNON C., SHAHN M., PIELER T., WARD D.C.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DDBJ databases.
DR EMBL: U67082; AAC61661.1; -.
DR PFAM: PF01352; KRAB; 1.
DR PFAM: PF00096; zf-C2H2; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 6.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 591 AA; 68328 MW; 047D7E62 CRC32;

Query Match      25.8%; Score 291; DB 11; Length 591;
Best Local Similarity 37.3%; Pred. No. 8.1e-21;
Matches 62; Conservative 28; Mismatches 52; Indels 24; Gaps

QY 36 LERHVVSDFSRPTSGKMNCDCVLGSCISFNVLWVKRSHTERPFQCNCQGASFQTQGN 95

```

OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 98056818.
 RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
 RT "Conservation of a master hematopoietic switch gene during vertebrate
 evolution: isolation and characterization of ikaros from teleost and
 amphibian species.";
 RL Eur. J. Immunol. 27:3049-3058(1997).
 CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER
 CC (DELTA-A ELEMENT) OF THE CD3-DELTA GENE.
 CC FUNCTIONS IN THE SPECIFICATION AND THE MATURATION OF THE T
 CC LYMPHOCYTE.
 CC ALSO INTERACTS WITH A CRITICAL CONTROL ELEMENT IN THE TDT
 CC (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) PROMOTER AS WELL AS WITH
 CC THE PROMOTERS FOR OTHER GENES EXPRESSED DURING EARLY STAGES OF B
 CC AND T CELL DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND SPLEEN.
 CC NO EXPRESSION IN MUSCLE.
 CC -!- SIMILARITY: TO D.MELANOGASTER HUNCHBACK.
 DR EMBL; U92202; AAB53435.1; -.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
 DR PFAM; PF00096; zf-C2H2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein.
 FT NON_TER 1 1
 FT ZN_FING 5 27 C2H2-TYPE.
 FT ZN_FING 33 55 C2H2-TYPE.
 FT ZN_FING 61 84 C2H2-TYPE.
 FT NON_TER 138 138
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 15701 MW; 2B786CAD CRC32;

 Query Match 48.6%; Score 547.5; DB 13; Length 138;
 Best Local Similarity 75.4%; Pred. No. 9.8e-47;
 Matches 104; Conservative 9; Mismatches 18; Indels 7; Gaps 1;

 QY 77 GERPFQCNQCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVKPK 136
 DB 1 GERPFQCNQCGASFTQKGNLLRHILKHTGERPFKCHLCNYACORRDALTGHLRTHSVKPK 60

 QY 137 YKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 189
 DB 61 HKCGYCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM-----GSEAL 189

 QY 190 VLDRLASNVAKRKSSMPQ 207
 DB 121 VLDRLASNVAKRKSSMPQ 138

 RESULT 6
 ID 042244 PRELIMINARY; PRT; 328 AA.
 AC 042244;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE IKAROS-RELATED TRANSCRIPTION FACTOR (FRAGMENT).
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 CC Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae;
 CC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TURPEN J., KELLEY C., MEAD P., ZON L.;
 RL Immunity 0:0-0(1997).
 DR EMBL; AF024439; AAB81280.1; -.
 DR PFAM; PF00096; zf-C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 2.
 KW Zinc-finger; Metal-binding; DNA-binding.
 FT NON_TER 1 1
 FT NON_TER 328 328

SQ SEQUENCE 328 AA; 37116 MW; CCC9E71E CRC32;

 Query Match 30.9%; Score 347.5; DB 13; Length 328;
 Best Local Similarity 55.3%; Pred. No. 1.2e-26;
 Matches 68; Conservative 13; Mismatches 27; Indels 15; Gaps 2;

 QY 102 LHTGEKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEHKRCRT 161
 DB 1 IRLGEKPFKPCFCYACRRRDALTGHLRTHSVGPKPHKNTCYGSRYSKORSLEEHKRCRN 60

 QY 162 FLQS-----TDPGDTASAEARHIKAEMG---SERALVLDRLASNVAKRKSSMP 206
 DB 61 YLQNVGMEAGQIVHHASPMDECKNPETMMNMSLMPFERPAVIERLASNMGRKKSSTP 120

 QY 207 QKF 209
 DB 121 QRF 123

 RESULT 7
 ID 062512 PRELIMINARY; PRT; 346 AA.
 AC 062512;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE ZINC FINGER PROTEIN (FRAGMENT).
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ICR; TISSUE=TESTIS;
 RA FUJIWARA Y.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ICR; TISSUE=TESTIS;
 RA FUJIWARA Y.;
 RT "Studies on zinc-finger protein genes expressed in mouse
 spermatogenesis.";
 RL Thesis (1991), Unknown Institution.
 DR EMBL; D10627; BAA01477.1; -.
 DR PFAM; PF00096; zf-C2H2; 12.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 12.
 KW Zinc-finger; Metal-binding; DNA-binding.
 FT NON_TER 1 1
 FT NON_TER 346 346
 SQ SEQUENCE 346 AA; 39665 MW; 06C78AAD CRC32;

 Query Match 26.5%; Score 298.5; DB 11; Length 346;
 Best Local Similarity 35.1%; Pred. No. 8.3e-22;
 Matches 59; Conservative 34; Mismatches 68; Indels 7; Gaps 3;

 QY 16 EPEIPYSYSREYNEYENIKLERHVVSFDSRPTSGRMNCVDCGLSCISFNVLVHKRSH 75
 DB 60 EKPEYECGCKGFAHYST--LRKHGSHHTGKPK---YECNECGKAFQAQSHLQNHKRT 113

 QY 76 TGERPFCNOCGASFTQKGNLLRHILKHTGEKPFKCHLCNYACORRDALTGHLRTHSVK 135
 DB 114 TGERPYECNECGKAFSAQSHSLQNHKRTHTGKPYECNECGKAFSAQSHSLQNHKRTHTG 173

 QY 136 YKCEFCGRSYKORSLEEHKRCRTFLQSTDPGDTASAEARHIKAEM 183
 DB 174 PYECNECGKAFSAQSHSLQNHKRTHTGKPYECNECGKAFSAQSHSLQNHKRTHTG 220

 RESULT 8
 ID 062514 PRELIMINARY; PRT; 594 AA.
 AC 062514;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE IKAROS.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OC Cyprinoidae; Cyprinidae; Rasbora; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA AMEMIYA C., KAWASAKI H.;
 RT "Characterization of zebrafish ikaros, a gene necessary for
 RT differentiation of the immune system.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF092175; AAC61763.1;
 DR PFAM: PF00096; zf-C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 537 AA; 58865 MW; 6AF8330F CRC32;

 Query Match 57.7%; Score 650; DB 13; Length 537;
 Best Local Similarity 55.7%; Pred. No. 3.5e-56;
 Matches 137; Conservative 17; Mismatches 38; Indels 54; Gaps 6;

 QY 3 DENVLSEPMGNAREPEI-----PYSREYNEYNKILRHVVSF 43
 DB 66 EENGLSCENGEAECAEDRLDLSGAKVNGSHAGPDKPAAYTAGGIRL----- 119

 QY 44 DSSRPTSGKMNCDVCLSCISFNVLVHKRSH-----GRPFQCNOCGASFTQ 92
 DB 119 -----PGLKLCIDIGVIGPNVLVHKRSHTEERKSVLEQQKGRPFQCNOCGASFTQ 173

 QY 93 KGNLRHKLHTGEKPKFCHLCNYACQRDALTGLRTHSVKPKCEFCGRSYKORSSL 152
 DB 174 KGNLRHKLHSGEKFCHLCNYACRRDALTGLRTHSVGPKHCAYCGRSYKORSSL 233

 QY 153 EEKHKRCRTFLOS-----TPDGTASAEARHIAEM-GSERALVLDRLASNVAKRKS 203
 DB 234 EEKHKCHNYLQCMGLQNSIYTVKEENSQNEQRE---DMPASERALVLDRIANNVAKRKS 290

 QY 204 SMPQKE 209
 DB 291 SMPQRF 296

 RESULT 3
 C92222 PRELIMINARY; PRT; 533 AA.
 AC Q92222;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE FOS PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR;
 RA HOMMA Y., KIYOSAWA H., MORI T., OGURI A., NIKAIKO T., KANAZAWA K.,
 RA TOJO M., TAKEDA J., TANNI Y., YOKOYA S., KAWABATA I., IKEDA H.,
 RA WANAKA A.;
 RT "Fos: a novel member of the Ikaros gene family expressed
 RT predominantly in the developing nervous system.";
 RL FEBS Lett. 0:0-0(1999).
 DR EMBL: AB017615; BAA36213.1;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 533 AA; 58167 MW; 2A0315AD CRC32;

 Query Match 54.8%; Score 617; DB 11; Length 533;

Best Local Similarity 53.6%; Pred. No. 6.1e-53;
 Matches 127; Conservative 23; Mismatches 33; Indels 54; Gaps 8;

 QY 1 ERDENVL-----KSEPM---GNAEEPEIPYSREYNEYNKILRHVVSFSSRPTSGK 52
 DB 68 DKDDSVIVEDLSLEPLGYCDGSGPEHPSP-----GGIRLDPNGK 105

 QY 53 MNCDCVCLSCISFNVLVHKRSHTEERPFQCNOCGASFTQKGNLRHKLHTGERKPKCH 112
 DB 106 LKCDVCGMVICGPNVLVHKRSHTEERPFQCNOCGASFTQKGNLRHKLHSGERPKCP 165

 QY 113 LONYACQRDALTGLRTHSV-----KPKCEFCGRSYKORSSLEEKRCRTFLO--S 165
 DB 166 FCNYACRRDALTGLRTHSVSSPTVGPKNYCNIGRSYKQOSTLEEKRCRHNYSLS 225

 QY 166 TD-----PGDTASAEARHK-----AEMGERALVLDRLASNVAKRKSMPQKE 209
 DB 226 TDAQALTQPGD-----EIDLEWVPSMLHPSTERTFIDRLANSLTKRKRSTPQKF 278

 RESULT 4
 P79751 PRELIMINARY; PRT; 417 AA.
 AC P79751;
 DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE IKAROS-LIKE.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
 OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GELLNER K., BRENNER S.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF056116; AAC34387.1;
 DR PFAM: PF00096; zf-C2H2; 5.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 417 AA; 45664 MW; 7856E12F CRC32;

 Query Match 51.4%; Score 578.5; DB 13; Length 417;
 Best Local Similarity 62.0%; Pred. No. 2.9e-49;
 Matches 103; Conservative 25; Mismatches 35; Indels 3; Gaps 1;

 QY 47 RPTSGKMNCDVCLSCISFNVLVHKRSHTEERPFQCNOCGASFTQKGNLRHKLHTGE 106
 DB 112 RLPNGKLOCEICGIVCTGPNVLVHKRSHTEERPFQCNOCGASFTQKGNLRHKLHSGE 171

 QY 107 KPFKCHLCNYACQRDALTGLRTHSVKPKCEFCGRSYKORSSLEEKRCRTFLOST 166
 DB 172 KPFKCPICNYACRRDALTGLRTHAVGKAFKCSYCSYKQOSTLEEDHQRCHSYLKR 231

 QY 167 DPGDTASAEARHKAE---MGSERALVLDRLASNVAKRKSMPQKE 209
 DB 232 QAAVVRQAAPGALNMDGMNOSNEKIQVDRDLANTIAKRKRTPQKF 277

 RESULT 5
 O13100 PRELIMINARY; PRT; 138 AA.
 AC O13100;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1)
 DE (FRAGMENT).
 GN IKAROS OR LYF-1.
 OS Xenopus laevis/gilli.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;

Search completed: November 6, 1999, 11:46:21
Job time: 6525 sec

```

FT  ZN_FING  268      290      C2H2-TYPE.
FT  ZN_FING  296      318      C2H2-TYPE.
FT  ZN_FING  324      346      C2H2-TYPE.
SQ  SEQUENCE  347 AA;  39855 MW;  F042FE59  CRC32;

Query Match      25.8%;  Score 290;  DB 1;  Length 347;
Best Local Similarity 34.8%;  Pred. No. 5.2e-19;
Matches 54;  Conservative 34;  Mismatches 61;  Indels 6;  Gaps 2;

QY  2  RDNVLKSEPMGNAEEPIPYSYSREYNEYENIKLERHVVSFDSRPTSGRMNCDVCLGS 61
    :  ::||| :  ||| :  ::||| :  :  :  :  :  :  :  :  :  :  :  :  :
Db  111 QSSGLLKQRIHTGKPYTCNVCDKHFIERSLTV--HQRHTTGEKP----YKCECGKA 164

QY  62  CISFNVLVVKRSHTGRRPPOCNCGASFTQGNLLRHILKHLTGEKPFKCHLQNYACORR 121
    :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :
Db  165 FSQSMNLTVHQRHTTGKPYTCNCGKAFRNSSLIQRHTIHTGKPYKCHDCGKAFSKN 224

QY  122 DALTGHLRTHSVEKPYKCFGRSYKORSSLEHK 156
    :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :
Db  225 SSLTQHRRIHTGKPYECMIGCKHFTGRSSLTVHQ 259
    :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :  ||| :

```

```

RESULT 13
OZF_BOVIN
ID OZF_BOVIN STANDARD; PRT; 292 AA.
AC Q28151;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN OZF.
CS ZNF146 OR OZF.
CS 90% TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.
AN [1]
RP SEQUENCE FROM N.A.
AC ALSSUE=LUNG;
AC MEDLINE; 96152243.
RA LE CHALONY C., APIOU F., PIBOUIN L., DUTRILLAUX B., GOUBIN G.;
RT "Constitutive amplification of a zinc finger protein gene in cattle.";
RL DNA CELL BIOL. 15:83-88(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE KRUEPEL SUBFAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/)
or send an email to license@lsb-sib.ch).
-----
CC
CC EMBL; X81804; E213761; ALT_INIT.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 10.
DR PFAM; PF00096; zf-C2H2; 10.
DR HSP; P25490; 12NM.
KW ZINC-FINGER; DNA-BINDING; METAL-BINDING; NUCLEAR PROTEIN; REPEAT.
FT DOMAIN 16 292 ZINC-FINGERS.
FT ZN_FING 16 38 C2H2-TYPE.
FT ZN_FING 44 66 C2H2-TYPE.
FT ZN_FING 72 94 C2H2-TYPE.
FT ZN_FING 100 122 C2H2-TYPE.
FT ZN_FING 128 150 C2H2-TYPE.
FT ZN_FING 156 178 C2H2-TYPE.
FT ZN_FING 184 206 C2H2-TYPE.
FT ZN_FING 212 234 C2H2-TYPE.
FT ZN_FING 240 262 C2H2-TYPE.
FT ZN_FING 268 292 C2H2-TYPE.
SQ SEQUENCE 292 AA; 33269 MW; 88B8A3DD CRC32.

```

```

Query Match          25.7%; Score 289; DB 1; Length 292;
Best Local Similarity 30.4%; Pred. No. 5.3e-19;
Matches 62; Conservative 45; Mismatches 67; Indels 30; Gaps 5;

Qy   4 ENVLKSEPMGNAEPELPYPSVSRNEYENIKLHRHVVSFDSSRP-----TSGKMN 54
      ||| : : : | : : : | : : : | : : : | : | | | |
Db   85 ENLTGHOKIHGEPFCKDCGKAFIQKSN--LIHQHRTHTGEKPFICKCGKTFSQSKN 142
      ||| : : : | : : : | : : : | : : : | : | | | |
Qy   55 -----CDVCGLSGISFNVLVMVKRSHSTGERPFCQNCOCGASFTQKGNNLRH 99
      ||| : : : | : : : | : : : | : : : | : | | | |
Db   143 LTHEKHITHGEKPKFKNCEGTAFQKKYLKHKQNIHGKPYECNECGKAFAQSRTSLIVH 202
      ||| : : : | : : : | : : : | : : : | : | | | |
Qy   100 IKLTGTEKPKFKCHLCNYACORRDALTGLRTHSVKPYKCFGCGRYSKQRSSLEEHKERC 159
      ::||::||::||:| : : ||::||::||:| : : ||::||:|
Db   203 VRIHSGDKPYECNVCGRAFGSQSSSLTVHRSHTGKPYGCGNECGKAFSQSFSTLALHL-RI 261
      ||| : : : | : : : | : : : | : : : | : | | | |
Qy   160 RTF---LQSTDPDGDTSAAEAARIHK 180
      | : : | : : | : : | : : | : : | : : | : : |
Db   262 HTGRKPYQCSCGKRAFQSKSHHIR 285
      ||| : : : | : : : | : : : | : : : | : | | | |

RESULT 14
Z141_HUMAN STANDARD; PRT; 474 AA.
ID Z141_HUMAN AC Q15928;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 141.
GN ZNF141.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAOGA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
RN PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RC [1]
RP SEQUENCE FROM N.A.
RX TISSUE=INSULINOMA;
RC MEDLINE; 94093541.
RA TOMMERUP N., AAGAARD L., LUND C.L., BOEL E., BAXENDALE S., BATES G.P.,
RA LEHRACH H., VISSING H.;
RT "A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate
RT gene for the Wolf-Hirschhorn (4p-) syndrome.";
RL HUM. MOL. GENET. 2:1571-1575(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION AS A
CC REPRESSOR. CANDIDATE GENE FOR WOLF-HIRSCHHORN (4p-) SYNDROME
CC (WHS).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL);
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY LOW EXPRESSION.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L15309; G347906; -.
DR MIM; 194648; -.
DR DR PROSITE; PS00028; ZINC_FINGER_C2H2; 10.
DR PFAM; PF00096; zf-C2H2; 10.
DR HSP; P08047; I5P2.
KW TRANSCRIPTION REGULATION; REPRESSOR; ZINC-FINGER; METAL-BINDING;
KW NUCLEAR PROTEIN.
FT DOMAIN 2 43 KRAB BOX ("A BOX").
FT DOMAIN 44 75 KRAB BOX ("B BOX").
FT DOMAIN 199 473 ZINC-FINGERS.
FT ZN_FING 199 221 C2H2-TYPE.
FT ZN_FING 227 249 C2H2-TYPE.
FT ZN_FING 255 277 C2H2-TYPE.
FT ZN_FING 283 305 C2H2-TYPE.

```

48 PMSGK-MNCDCVGLSCISFNVLVHRSHTGPRFQCNOCGASFTQKGNLLRHKLHTGE 106
 208 PMSGKPEYECNHCCKVLSHKGQGLDQRTTGKPYECNCGIAFSQKSLVYHQRTHTGE 267
 107 KPFKCHLNCVACORRDLGHLRTSHVEKPYKCEFCGRSYKORSSLEH----- 156
 268 KPYECQCKKANGHKLHDLRIHTGEPKYNCEGKTFRRSSNLMQHLRSHTGEKPYE 327
 156 -KERCRTFLOSTDPGTASAEARHKAEMG 184
 328 CRECGKSFYN-----SSLTEHVRHTG 350

RESULT 11
 KIDL_MOUSE
 ID KIDL_MOUSE STANDARD; PRT; 572 AA.
 AC Q61751;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).
 GN TCF17 OR KID1.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE-LENS;
 RA BRADY J.P., DUNCAN M.K., WAWROUSEK E.F., PIATIGORSKY J.;
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: IT MAY PLAY A ROLE IN RENAL DEVELOPMENT AND MAY
 CC ALSO BE INVOLVED IN THE REPAIR OF THE KIDNEY AFTER ISCHEMIA-
 CC REPERFUSION OR FOLIC ACID ADMINISTRATION.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
 CC

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

EMBL; L77247; G1256362; -
 DR MGD; MGI:103172; TCF17.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 13.
 DR PFAM; PF00096; zf-C2H2; 13.
 DR HSSP; P08047; ISP2.
 KW ZINC-FINGER; TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;
 NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; METAL-BINDING.
 FT DOMAIN 12 53
 FT KRAB BOX ("A BOX").
 FT KRAB BOX ("B BOX").
 FT ZINC-FINGERS.
 FT DOMAIN 184 567
 FT ZN_FING 184 206
 FT ZN_FING 212 234
 FT C2H2-TYPE.
 FT ZN_FING 240 262
 FT C2H2-TYPE.
 FT ZN_FING 268 290
 FT C2H2-TYPE.
 FT ZN_FING 321 343
 FT C2H2-TYPE.
 FT ZN_FING 349 371
 FT C2H2-TYPE.
 FT ZN_FING 377 399
 FT C2H2-TYPE.
 FT ZN_FING 405 427
 FT C2H2-TYPE.
 FT ZN_FING 431 455
 FT C2H2-TYPE.
 FT ZN_FING 461 483
 FT C2H2-TYPE.
 FT ZN_FING 489 511
 FT C2H2-TYPE.
 FT ZN_FING 517 539
 FT C2H2-TYPE.
 FT ZN_FING 545 567
 FT C2H2-TYPE.
 SQ SEQUENCE 572 AA; 65720 MW; 4F7FBBDC CRC32;

Query Match 26.0%; Score 293; DB 1; Length 572;

Best Local Similarity 33.7%; Pred. No. 4.8e-19;
 Matches 61; Conservative 36; Mismatches 60; Indels 24; Gaps 5;
 QY 19 EIPYSYSREYENYENI-KLERHVVVSFDSSRPTSGMKMCDVGLSCISFNVLVHRSHTG 77
 DB 374 EKPFCSECGKGFISRLNHRHIIHTGEK-----FYCNCEGKALSSHTLIHERIHTG 429
 QY 78 ERPFQCNOCGASFTQKGNLLRHKLHTGEPKPFKCHLNCVACORRDLGHLRTSHVEKPY 137
 DB 430 EKPCCKVCGKAFQSSALIQHQMHTGTERPYKCNCEGKTFRCNSSLNHSNORTHTGKPY 489
 QY 138 KCEFCGRSYKORSSLEHKE-----RC-----RTFLOSTDPGTASAEARHKAEMGS 185
 DB 490 RCQEGMSFGOSSALIQHRIHTGEPKFCNCTGKTRQS-----SSRIAHQRIHTGE 542
 QY 186 E 186
 DB 543 K 543

RESULT 12
 KR2_MOUSE
 ID KR2_MOUSE STANDARD; PRT; 347 AA.
 AC P08043;
 DT 01-AUG-1988 (REL. 08, CREATED)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE MKR2 PROTEIN (ZINC FINGER PROTEIN 2).
 GN ZFP2 OR MKR2.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 88312581.
 RA CHOWDHURY K., DRESSLER G., DREIER G., DEUTSCH U., GRUSS P.;
 RT "The primary structure of the murine multifinger gene mkr2 and its
 RT specific expression in developing and adult neurons."
 RL EMBO J. 7:1345-1353(1988).
 RN
 RP SEQUENCE OF 75-347 FROM N.A.
 RX MEDLINE: 87131089.
 RA CHOWDHURY K., DEUTSCH U., GRUSS P.;
 RT "A multigene family encoding several 'finger' structures is present
 RT and differentially active in mammalian genomes."
 RL CELL 48:771-778(1987).
 CC -!- FUNCTION: CAN BE A REGULATORY FACTOR REQUIRED FOR THE
 CC DIFFERENTIATION AND/OR PHENOTYPIC MAINTENANCE OF NEURONS.
 CC

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

EMBL; Y08850; G52811; -
 DR EMBL; M15709; G193352; -
 DR PIR; S00549; S00549.
 DR PIR; B26595; B26595.
 DR MGD; MGI:99167; ZFP2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 9.
 DR PFAM; PF00096; zf-C2H2; 9.
 DR HSSP; P08047; ISP2.
 KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; REPEAT.
 FT DOMAIN 100 346
 FT ZN_FING 100 122
 FT C2H2-TYPE.
 FT ZN_FING 128 150
 FT C2H2-TYPE.
 FT ZN_FING 156 178
 FT C2H2-TYPE.
 FT ZN_FING 184 206
 FT C2H2-TYPE.
 FT ZN_FING 212 234
 FT C2H2-TYPE.
 FT ZN_FING 240 262
 FT C2H2-TYPE.


```

FT ZN_FING 107 129 C2H2-TYPE.
FT ZN_FING 135 157 C2H2-TYPE.
FT ZN_FING 163 185 C2H2-TYPE.
FT ZN_FING 191 213 C2H2-TYPE.
FT ZN_FING 236 258 C2H2-TYPE.
FT ZN_FING 264 286 C2H2-TYPE.
FT ZN_FING 292 314 C2H2-TYPE.
FT ZN_FING 320 342 C2H2-TYPE.
FT ZN_FING 348 370 C2H2-TYPE.
FT ZN_FING 376 398 C2H2-TYPE.
FT ZN_FING 404 426 C2H2-TYPE.
FT ZN_FING 432 454 C2H2-TYPE.
FT ZN_FING 460 482 C2H2-TYPE.
SQ SEQUENCE 485 AA; 54463 MW; 908B33F1 CRC32;

Query Match 26.4%; Score 297; DB 1; Length 485;
Best Local Similarity 41.0%; Pred. No. 1.8e-19;
Matches 57; Conservative 22; Mismatches 50; Indels 10; Gaps 2;

QY 18 PEIPYSREYNEVENIKLERHHVVSFSSRPTSGKMKNCVDCGLSCISFNLVLMVHKRSHTG 77
DDB 351 PCGQRFSS-----QRLKLRHFQHTGEXP-----YHCGEGLGTQVSRLTEHQRIHTG 400
QY 78 ERPFCNOCGASFTOKNLRLHRLKLTGKPKFKCHLCNYACQRRDALTGHLRTHSVEKPY 137
DDB 401 ERPFACPCGSGFROHANLTQHRRIHTGERPYACPCGKAFQRPRTLQHLRTHRREKPF 460
QY 138 KCEFCGRSYKQRSSLEEHK 156
DDB 461 ACQDCGRRFHQSTKLIOHQ 479

RESULT 9
ZN91_HUMAN
ID ZN91_HUMAN STANDARD; PRT; 1191 AA.
AC Q05481;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 91 (ZINC FINGER PROTEIN HTF10) (HPF7).
GN ZNF91.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 9323677.
RA BELLEFROID E.J., MARINE J.C., RIED T., LECOCQ P.J., RIVIERE M.,
RA AMEMIYA C.T., PONCELET D.A., COULIE P.G., DE JONG P.J.,
RA SZPIRER C., WARD D.C., MARTIAL J.A.;
RT "Clustered organization of homologous KRAB zinc-finger genes with
RT enhanced expression in human T lymphoid cells.";
RL EMBO J. 12:1363-1374 (1993).
RN [2]
RN SEQUENCE OF 15-204 FROM N.A.
RX MEDLINE; 91219421.
RA BELLEFROID E.J., PONCELET D.A., LECOCQ P.J., REVELANT O.,
RA MARTIAL J.A.;
RT "The evolutionarily conserved Kruppel-associated box domain defines a
RT subfamily of eukaryotic multifingered proteins.";
RT PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612 (1991).
CC CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC CC FINGER PROTEINS.
CC CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
DR PDB: 1ZNF: 15-OCT-91.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 35.
DR PFAM: PF00096; zf-C2H2; 36.
KW ZINC-FINGER; METAL-BINDING; RNA-BINDING; REPEAT; 3D-STRUCTURE;
KW PHOSPHORYLATION.
FT DOMAIN 1 26 KRAB BOX ("A BOX").
FT DOMAIN 27 58 KRAB BOX ("B BOX").
FT DOMAIN 108 298 ZINC-FINGERS I.
FT DOMAIN 326 488 ZINC-FINGERS II.
FT DOMAIN 503 721 ZINC-FINGERS III.
FT DOMAIN 750 940 ZINC-FINGERS IV.
FT DOMAIN 988 1066 ZINC-FINGERS V.
FT DOMAIN 1136 1298 ZINC-FINGERS VI.
FT ZN_FING 108 130 C2H2-TYPE.
FT ZN_FING 136 158 ZN_FING 136 158
FT ZN_FING 164 186 C2H2-TYPE.
FT ZN_FING 192 214 C2H2-TYPE.
FT ZN_FING 220 242 C2H2-TYPE.
FT ZN_FING 248 270 C2H2-TYPE.
FT ZN_FING 276 298 C2H2-TYPE.
FT ZN_FING 326 348 C2H2-TYPE.
FT ZN_FING 354 376 C2H2-TYPE.
FT ZN_FING 382 404 C2H2-TYPE.
FT ZN_FING 410 432 C2H2-TYPE.
FT ZN_FING 438 460 C2H2-TYPE.
FT ZN_FING 466 488 C2H2-TYPE.
FT ZN_FING 503 525 C2H2-TYPE.
FT ZN_FING 531 553 C2H2-TYPE.
FT ZN_FING 559 587 C2H2-TYPE.
FT ZN_FING 587 609 C2H2-TYPE.
FT ZN_FING 615 637 C2H2-TYPE.
FT ZN_FING 643 665 C2H2-TYPE.
FT ZN_FING 671 693 C2H2-TYPE.
FT ZN_FING 699 721 C2H2-TYPE.
FT ZN_FING 750 772 C2H2-TYPE.
FT ZN_FING 778 800 C2H2-TYPE.
FT ZN_FING 806 828 C2H2-TYPE.
FT ZN_FING 834 856 C2H2-TYPE.
FT ZN_FING 862 884 C2H2-TYPE.
FT ZN_FING 890 912 C2H2-TYPE.
FT ZN_FING 918 940 C2H2-TYPE.
FT ZN_FING 988 1010 C2H2-TYPE.
FT ZN_FING 1016 1038 C2H2-TYPE.
FT ZN_FING 1044 1066 C2H2-TYPE.
FT ZN_FING 1136 1158 C2H2-TYPE.
FT ZN_FING 1164 1186 C2H2-TYPE.
FT ZN_FING 1192 1214 C2H2-TYPE.
FT ZN_FING 1220 1242 C2H2-TYPE.
FT ZN_FING 1248 1270 C2H2-TYPE.
FT ZN_FING 1276 1298 C2H2-TYPE.
FT STRAND 1045 1045
FT STRAND 1052 1052
FT HELIX 1056 1066
FT TURN 1066 1066
SQ SEQUENCE 1350 AA; 155804 MW; 10CA7C8C CRC32;

Query Match 28.2%; Score 317.5; DB 1; Length 1350;
Best Local Similarity 36.9%; Pred. No. 7.6e-21;
Matches 66; Conservative 33; Mismatches 59; Indels 21; Gaps 4;

QY 4 ENVLKSEPMGNAREEIPYYSREYNEVENIKLRHVVSFDSSRPTSGKMN---CDVCG 60
Db 308 DSVGTDPLSSQVASSPYSCSKRKTFRWK-----SFLNHQOHTSRKPYLCSHCNK 361
QY 61 SCISFNLVWVKRSTGTERPPFCNCGSFTQKGNLLRHLKLTGCKEFPKCHLNYACQR 120
Db 362 GFQNSDLVKHRTHTGTERPPYOCACHGFTQKSDLVKHLRTHRTGCKEFPKCHDCKFTE 421
QY 121 RDALTGLRTHRSVEPKCEFGCRKYKORSSEEHKE-----RC-----RTFLQSTD 167
Db 422 RSALAKHQRTHGTERPPYKSCDCKEFTQRSNLLHQRHTGTERPPYKCYLDCRTFIQNSD 480
```

```
RESULT 7
ZN43_HUMAN
ID ZN43_HUMAN STANDARD; PRT; 803 AA.
AC P28160;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 43 (ZINC PROTEIN HTF6).
GN ZNF43.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-CELL;
RX MEDLINE: 91279444.
RA LOVERING R., TROWSDALE J.;
RT "A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell lines.";
RL NUCLEIC ACIDS RES. 19:2921-2927(1991).
[2]
RP SEQUENCE OF 38-190 FROM N.A.
RX MEDLINE: 91219421.
RA BELLEFROID E.J., PONCELET D.A., LECOCQ P.J., REVELANT O., MARTIAL J.A.;
RT "The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins.";
RL PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612(1991).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: T AND B CELL LINES.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59244; G38032;
DR EMBL: M61869; G184450;
DR PIR: S26823; S26823.
DR PIR: D39384; D39384.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 19.
DR PFAM: PF00096; zf-C2H2; 21.
DR HSSP: P25490; 1ZNF.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN.
FT DOMAIN 1 37 KRAB BOX ("A BOX").
FT DOMAIN 38 69 KRAB BOX ("B BOX").
FT DOMAIN 168 778 22 C2H2-TYPE ZINC-FINGERS.
FT ZN_FING 168 190 C2H2-TYPE.
FT ZN_FING 196 218 C2H2-TYPE (DEGENERATE).
FT ZN_FING 224 246 C2H2-TYPE (DEGENERATE).
FT ZN_FING 252 274 C2H2-TYPE (DEGENERATE).
FT ZN_FING 280 302 C2H2-TYPE.
FT ZN_FING 308 330 C2H2-TYPE.
FT ZN_FING 336 358 C2H2-TYPE.
FT ZN_FING 364 386 C2H2-TYPE.
FT ZN_FING 392 414 C2H2-TYPE.
FT ZN_FING 420 442 C2H2-TYPE.
FT ZN_FING 448 470 C2H2-TYPE.
FT ZN_FING 476 498 C2H2-TYPE.
FT ZN_FING 504 526 C2H2-TYPE.
FT ZN_FING 532 554 C2H2-TYPE.
FT ZN_FING 560 582 C2H2-TYPE.
FT ZN_FING 610 638 C2H2-TYPE.
FT ZN_FING 616 638 C2H2-TYPE.
FT ZN_FING 644 666 C2H2-TYPE.
FT ZN_FING 672 694 C2H2-TYPE.
```


CC AND AT DAY 5-6 IN THE EMBRYO PROPER.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U92201; G2062742; -;
DR EMBL; U92200; G2062740; -;
DR EMBL; U92198; G2062736; -;
DR EMBL; U92199; G2062738; -;
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
DR PFAM; PF00096; zf-C2H2; 5.
KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 125 203
FT DOMAIN 468 520
FT ZN_FING 125 147
FT ZN_FING 153 175
FT ZN_FING 181 203
FT ZN_FING 209 232
FT ZN_FING 468 490
FT ZN_FING 496 520
FT VARSPLIC 55 148
FT VARSPLIC 55 148
FT VARSPLIC 149 291
FT VARSPLIC 206 246
FT VARSPLIC 206 291
FT VARSPLIC 247 291
SQ SEQUENCE 522 AA; 57657 MW; F39835EE CRC32;

Query Match 59.9%; Score 675; DB 1; Length 522;
Best Local Similarity 58.1%; Pred. No. 2.8e-53;
Matches 136; Conservative 23; Mismatches 39; Indels 36; Gaps 4;

QY 3 DENLVKSEPMGNAEPEIYSYSREYNEVENIKL-----ERHVSFDSRPTS--- 51
DB 66 EENGLAIDMMNGEET-----ECAEDLVLDASAKYNGHAGDPSKGYSSAG 116

QY 51 -----GKNMCDVGLSCISFNVLVHKRSHTGERPFCQNCQCGASFTQGNLRLHKLHT 104
DB 117 GIRLPNGKLKDCIGVICGPNVLVHKRSHTGERPFCQNCQCGASFTQGNLRLHKLHS 176

QY 105 GEKPKFCHLCNACORRDALTGLRTHSVKPKYKCFECGRSYKQSSLEEHKERCRTFLQ 164
DB 177 GEKPKFCHLCNACORRDALSGLRTHSVGKPKKCAICGRSYKQSSLEEHKERCRTFLQ 236

QY 165 -----STDGPDTSAAEARIHKAEMGSEALVLDRLASNVAKRKSMPQKF 209
DB 237 CMGLONSIYTVVKEESNONEQREDLSQMGSKRALVLDRLANNVAKRKSMPQKF 290

RESULT 4
ID IKAR MOUSE STANDARD; PRT: 517 AA.
AC Q03267; Q64044; Q64045; Q64051;
DT 01-OCT-1993 (REL. 27, CREATED)
DI 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).
GN IKAROS OR LYFI.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 5).
RC TISSUE=EMBRYO;
RX MEDLINE; 93068267.

RA GEORGOPOULOS K., MOORE D.D., DERFLER B.;
RT "Ikaros, an early lymphoid-specific transcription factor and a
RL putative mediator for T cell commitment.";
RN SCIENCE 258:808-812(1992).
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE; 95021239.
RA HAM K., ERNST P., LO K., KIM G.S., TURCK C., SMALE S.T.;
RT "The lymphoid transcription factor Lyf-1 is encoded by specific,
RL alternatively spliced mRNAs derived from the Ikaros gene.";
CC MOL. CELL. BIOL. 14:7111-7123(1994).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED
CC DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN, ISOFORMS I-VI, ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT
CC OF ISOFORM VI.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN T-CELLS AND THEIR
CC PROGENITORS, AND ALSO IN B-CELLS.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03547; G198287; -;
DR EMBL; S74517; G807149; ALT_SEQ.
DR EMBL; S74518; G807151; -;
DR EMBL; S74708; G807153; -;
DR MGD; MGI:965335; IKAROS.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
DR PFAM; PF00096; zf-C2H2; 4.
DR HSSP; P08047; LSP2.
DR TRANSFAC; T01470; -;
KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 117 223
FT DOMAIN 457 479
FT ZN_FING 117 139
FT ZN_FING 144 166
FT ZN_FING 172 194
FT ZN_FING 200 223
FT ZN_FING 457 479
FT VARSPLIC 53 53
FT VARSPLIC 54 140
FT VARSPLIC 54 282
FT VARSPLIC 141 282
FT VARSPLIC 234 235
FT CONFLICT 234 235
FT CONFLICT 480 482
SQ SEQUENCE 517 AA; 57336 MW; D18D2D97 CRC32;
MISSING (IN ISOFORM V).
MISSING (IN ISOFORMS I AND II).
MISSING (IN ISOFORMS III AND IV).
VC -> MV (IN REF. 2).
MISSING (IN REF. 2).
M -> VAYGADGRDFHAIISDRGM (IN ISOFORMS
II AND IV).
Query Match 58.8%; Score 662.5; DB 1; Length 517;
Best Local Similarity 73.7%; Pred. No. 3.7e-52;
Matches 126; Conservative 15; Mismatches 21; Indels 9; Gaps 3;

QY 47 RPTSGKMNCDVGLSCISFNVLVHKRSHTGERPFCQNCQCGASFTQGNLRLHKLHTGE 106
DB 111 RLPNGKLKDCIGVICGPNVLVHKRSHT-ERPFCQNCQCGASFTQGNLRLHKLHSGE 169

QY 107 KPFKCHLCNACORRDALTGLRTHSVKPKYKCFECGRSYKQSSLEEHKERCRTFLQST 166
DB 170 KPFKCHLCNACORRDALTGLRTHSVGKPKKCAICGRSYKQSSLEEHKERCRTFLQST 229

FT VARSPLIC 10 53 MISSING (IN IK4).
 FT VARSPLIC 54 140 MISSING (IN IK2).
 FT VARSPLIC 54 283 MISSING (IN IK6).
 FT VARSPLIC 197 283 MISSING (IN IK3 AND IK4).
 FT VARSPLIC 141 283 MISSING (IN IK5).
 FT CONFLICT 11 12 QV -> FS (IN REF. 2).
 FT CONFLICT 214 214 S -> T (IN REF. 2).
 FT CONFLICT 245 245 N -> K (IN REF. 2).
 FT CONFLICT 296 296 MISSING (IN REF. 2).
 FT CONFLICT 298 298 S -> T (IN REF. 2).
 FT CONFLICT 352 352 KPLA -> RRS (IN REF. 2).
 FT CONFLICT 372 372 N -> Y (IN REF. 2).
 FT CONFLICT 420 426 PHARNGL -> RRAQV (IN REF. 2).
 SQ SEQUENCE 519 AA; 57528 MW; 5A97272D CRC32;

Query Match 60.8%; Score 684.5; DB 1; Length 519;
 Best Local Similarity 74.4%; Pred. No. 4e-54; 20; Indels 9; Gaps 2;
 Matches 128; Conservative 15; Mismatches 15;

QY 47 RPTSGKMCNDVGLSCISFNVLVHVKRSHTGERPFCQNCQASFTQKGNLLRHKLHTGE 106
 DB 111 RLPNGKLCDCIGIICGNPVLVHVKRSHTGERPFCQNCQASFTQKGNLLRHKLHSGE 170
 QY 107 KPFKCHLCYACORDALTGHLRTHSVKPYKCFGCGRSYKORSLEEHKRCRTFLQST 166
 DB 171 KPFKCHLCYACRRDALTGHLRTHSVGKPHKCGYCGRSYKORSLEEHKRCRCHYLESM 230
 QY 167 D-PG-----DTASAEARHKAEMGSRALVLDRLASNAVAKRKSMPQKF 209
 DB 231 GLPGTLVPYKETEHNSEMAEDLCKIGSERSLVLDRLASNAVAKRKSMPQKF 282

RESULT 2
 IKAR_CHICK
 ID IKAR_CHICK STANDARD; PRT; 518 AA.
 AC 042410;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DE DNA-BINDING PROTEIN IKAROS.
 GN IKAROS OR IK.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
 RN [1]
 RP "Avian Ikaros gene is expressed early in embryogenesis."
 RL EUR. J. IMMUNOL. 27:1853-1857(1997).
 CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC HEMATOPOIETIC ORGANS SUCH AS THE BURSA OF FABRICIUS, THYMUS AND SPLEEN. IN THE ADULT, EXPRESSED IN SPLEEN, THYMUS, BURSA AND PERIPHERAL BLOOD LEUKOCYTES.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYO FROM EMBRYONIC DAY 2 ONWARDS.
 CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y11833; E309211; -
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
 DR PFAM; PF00096; zf-C2H2; 5.
 KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
 FT DOMAIN 117 224 ZINC-FINGERS I.
 FT ZN_FING 461 513 ZINC-FINGERS II.
 FT ZN_FING 117 139 C2H2-TYPE.
 FT ZN_FING 145 167 C2H2-TYPE.
 FT ZN_FING 173 195 C2H2-TYPE.
 FT ZN_FING 201 224 C2H2-TYPE.
 FT ZN_FING 461 483 C2H2-TYPE.
 FT ZN_FING 489 513 C2H2-TYPE.
 SQ SEQUENCE 518 AA; 57586 MW; AB4DF123 CRC32;

Query Match 60.1%; Score 676.5; DB 1; Length 518;
 Best Local Similarity 73.3%; Pred. No. 2.1e-53;
 Matches 126; Conservative 14; Mismatches 23; Indels 9; Gaps 1;

QY 47 RPTSGKMCNDVGLSCISFNVLVHVKRSHTGERPFCQNCQASFTQKGNLLRHKLHTGE 106
 DB 111 RLPNGKLCDCIGIICGNPVLVHVKRSHTGERPFCQNCQASFTQKGNLLRHKLHSGE 170
 QY 107 KPFKCHLCYACORDALTGHLRTHSVKPYKCFGCGRSYKORSLEEHKRCRTFLQ-- 165
 DB -171 KPFKCHLCYACRRDALTGHLRTHSVGKPHKCGYCGRSYKORSLEEHKRCRCHNYLOTM 230
 QY 165 -----STDPGDTASAEARHKAEMGSRALVLDRLASNAVAKRKSMPQKF 209
 DB 231 SISSNLYSVKEETNQSEMAEDLCKIGSERSLVLDRLASNAVAKRKSMPQKF 282

RESULT 3
 IKAR_ONCMY
 ID IKAR_ONCMY STANDARD; PRT; 522 AA.
 AC 013089; 013088; 013086; 013087;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE DNA-BINDING PROTEIN IKAROS.
 GN IKAROS.
 OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIARDNERI).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;
 CC SALMONIDAE; ONCORHYNCHUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SHASTA: TISSUE-THYMOCYTES;
 RX MEDLINE: 98056818
 RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
 RT "Conservation of a master hematopoietic switch gene during vertebrate evolution: isolation and characterization of Ikaros from teleost and amphibian species."
 RL EUR. J. IMMUNOL. 27:3049-3058(1997).
 CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- ALTERNATIVE PRODUCTS: EIGHT FORMS OF THE PROTEIN, IK1-1K8; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF IK1.
 CC -!- TISSUE SPECIFICITY: EXPRESSION MAINLY LIMITED TO THYMUS, SPLEEN, AND PRONEPHROS. VERY LOW EXPRESSION IN LIVER. NO EXPRESSION IN TESTIS, BRAIN, EYE AND MUSCLE.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT DAY 3-4 IN THE YOLK SAC

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 11:46:20 ; Search time 13.55 Seconds
(without alignments)
436.020 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 1226
Sequence: 1 ERDENLKGSEPMGNAEPEI.....VLDRLASNAVKRSMPOKE 209

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	684.5	60.8	519	1 IKAR_HUMAN	Q13422 homo sapien
2	676.5	60.1	518	1 IKAR_CHICK	Q42410 gallus gall
3	675	59.9	522	1 IKAR_ONCMY	Q13089 oncorhynchu
4	662.5	58.8	517	1 IKAR_MOUSE	Q03267 mus musculu
5	618	54.9	526	1 HELI_MOUSE	P81183 mus musculu
6	317.5	28.2	1350	1 XFIN_XENLA	P08045 xenopus lae
7	299	26.6	803	1 ZN43_HUMAN	P28160 homo sapien
8	297	26.4	485	1 ZN42_HUMAN	P28698 homo sapien
9	296	26.3	1191	1 ZN91_HUMAN	Q05481 homo sapien
10	294	26.1	553	1 ZF37_MOUSE	P17141 mus musculu
11	293	26.0	572	1 KIDL_MOUSE	O61751 mus musculu
12	290	25.8	347	1 KR2_MOUSE	P08043 mus musculu
13	289	25.7	292	1 OZF_BOVIN	Q28131 bos taurus
14	289	25.7	474	1 Z141_HUMAN	Q15938 homo sapien
15	288	25.6	292	1 OZF_HUMAN	Q15072 homo sapien
16	288	25.6	629	1 Z11B_HUMAN	Q06732 homo sapien
17	288	25.6	629	1 Z195_HUMAN	O14628 homo sapien
18	288	25.6	595	1 ZN85_HUMAN	Q03923 homo sapien
19	287.5	25.5	439	1 Z028_XENLA	P18747 xenopus lae
20	285.5	25.4	614	1 ZF29_MOUSE	Q07230 mus musculu
21	285	25.3	506	1 Z157_HUMAN	P51786 homo sapien
22	285	25.3	848	1 Z33A_HUMAN	Q06730 homo sapien
23	284.5	25.3	626	1 Z189_HUMAN	O75820 homo sapien
24	282.5	25.1	510	1 CF2_DROME	P20385 drosophila
25	282.5	25.1	580	1 ZF35_MOUSE	P15620 mus musculu
26	282.5	25.1	428	1 ZN83_HUMAN	P51522 homo sapien
27	282	25.0	393	1 ML24_MOUSE	Q03309 mus musculu
28	282	25.0	336	1 ZG57_XENLA	P18729 xenopus lae
29	281.5	25.0	469	1 Z135_HUMAN	P52742 homo sapien
30	280.5	24.9	589	1 Z132_HUMAN	P52740 homo sapien
31	280	24.9	321	1 Z177_HUMAN	Q13360 homo sapien
32	280	24.9	543	1 ZN08_HUMAN	P17098 homo sapien
33	279	24.8	576	1 KIDL_RAT	Q02975 rattus norv
34	279	24.8	247	1 Z020_XENLA	P18744 xenopus lae
35	278.5	24.7	428	1 ZF26_MOUSE	P10076 mus musculu
36	278.5	24.7	229	1 ZN22_HUMAN	P17026 homo sapien
37	278	24.7	636	1 ZF90_MOUSE	Q61967 mus musculu
38	278	24.7	682	1 ZN45_HUMAN	Q02386 homo sapien
39	277	24.6	519	1 ZN35_HUMAN	P13682 homo sapien
40	276.5	24.6	726	1 Z184_HUMAN	Q99676 homo sapien
41	276.5	24.6	242	1 Z37A_HUMAN	P17032 homo sapien
42	275	24.4	453	1 Z06_XENLA	P18749 xenopus lae
43	274	24.3	586	1 ZN07_HUMAN	P17097 homo sapien

ALIGNMENTS

RESULT 1	IKAR_HUMAN	169	1	MFGL_MOUSE	P16372 mus musculu
44	273.5	24.3	407	1 OZFE_RAT	Q62981 rattus norv
45	273	24.2			
ID	IKAR_HUMAN	STANDARD;	PRT;	519 AA.	
AC	Q13422; 000598;				
DT	15-DEC-1998 (REL. 37, CREATED)				
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)				
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)				
DE	DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).				
GN	IKAROS OR IK1 OR OR LYF1.				
OS	HOMO SAPIENS (HUMAN)				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.				
RP	TISSUE=BONE MARROW;				
RC	MEDLINE; 96252222.				
RA	NIETFIELD W., MEYERHANS A.;				
RT	"Cloning and sequencing of hik-1, a cDNA encoding a human homologue of mouse Ikaro/Lyf-1.";				
RT	IMMUNOL. LETT. 49:139-141(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 96132884.				
RA	MOLNAR A., WU P., LARGESPADA D.A., VORTKAMP A., SCHERER S., COPELAND N.G., JENKINS N.A., BRUNS G., GEORGIOPOULOS K.;				
RT	"The Ikaro gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse.";				
RL	J. IMMUNOL. 156:585-592(1996).				
CC	-!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED. DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.				
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.				
CC	-!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMUS, SPLEEN AND PERIPHERAL BLOOD LEUCOCYTES AND LYMPH NODES. LOWER EXPRESSION IN BONE MARROW AND SMALL INTESTINE.				
CC	-!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN: ISOFORMS IK1-IK6; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF IK1.				
CC	-!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; U40462; G1289371; -				
DR	EMBL; S80876; G1911483; -				
DR	MIM; 603023; -				
DR	PROSITE; PS00028; ZINC_FINGER_C2H2; 5.				
DR	PFAM; PF00096; zf-C2H2; 5.				
KW	TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.				
FT	DOMAIN 117 224				
FT	ZINC-FINGERS I.				
FT	ZN_FING 117 139				
FT	C2H2-TYPE.				
FT	ZN_FING 145 167				
FT	C2H2-TYPE.				
FT	ZN_FING 173 195				
FT	C2H2-TYPE.				
FT	ZN_FING 201 224				
FT	C2H2-TYPE.				
FT	ZN_FING 462 484				
FT	C2H2-TYPE.				
FT	ZN_FING 490 514				
FT	C2H2-TYPE.				

RESULT 14

transcription repressor zinc finger protein 85 - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G02075
R:Poncellet, D.A.
submitted to the EMBL Data Library, September 1995
A:Reference number: G09169
A:Accession: G02075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-595 <PON>
A:Cross-references: EMBL:U35376; NID:gl017721; PID:gl017722
C:Genetics:
A:Gene: GDB:2NF85
A:Cross-references: GDB:132279
A:Map position: 19p12-19p12

```

Query Match      25.5%; Score 288; DB 2; Length 595;
Best Local Similarity 32.3%; Pred. No. 8,1e-18;
Matches 61; Conservative 25; Mismatches 41; Indels 62; Gaps 7;

QY 24 YSREYNEVENIKLERHVVSFDSSRPTSGKMNCDVCG-----LSCI----- 64
      | : : : | | : : : | | : | : | :
DB 151 YVKVAHKFSN -SNRHEIRHTKKP-----FKCTCGSGFGMISCLTHERSHTRVNFYKC 204
      | : : : | | : : : | | : | : | :
QY 64 -----SPN-----VLMVHKRSHTGERPFOCNQCGASFTQGNLLRHKIKLHTGKPKFKCHLCN 115
      | : | | | | | | | | | | | | | | | | | | | | |
DB 205 EECGKAFNWSSTLTKHKRIHTGEKPYKCECGKAFNQSNNLIKKIKHTGKPYKCEBGC 264
      | : | | | | | | | | | | | | | | | | | | | | |
QY 116 -----YACQ-----RDATGLHTRHSVEKPYKCECGRSYK 147
      | : | | | | | | | | | | | | | | | | | | | | |
DB 265 KAFNRFTLTHKIIHTGEKPYKCEGKAFNRSTLTHRKHITGEKPYKCEGKAFFK 324
      | : | | | | | | | | | | | | | | | | | | | | |
QY 148 QRSLEEKK 156
      | : | | |
DB 325 OSSNLTTHK 333
      | : | | |

```

RESULT 15

finger protein (clone XlCOP28) - African clawed frog (fragment)
 S06556
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990.#text_change 31-Dec-1993
 C:Accession: S06556
 R:Niefield, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Knoechel, J. Mol. Biol. 208, 639-659, 1989
 A:Title: Second-order repeats in Xenopus laevis finger proteins.
 A:Reference number: S05632; MUID:90040698
 #:Accession: S06556
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-439 <NIE>
 C:Keywds: DNA binding; zinc finger

```

Query Match          25.5%; Score 287.5; DB 2; Length 439;
Best Local Similarity 36.8%; pred. No. 6.6e-18;
Matches 64; Conservative 28; Mismatches 54; Indels 19; Gaps 5;

QY 36 LERVVVSFDSRPTSGKMNCDVCGLSGTSFNVLMVKRSHHTGERPFCNOCGASFQKGN 95
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 LVEHQRTHTGRP-----FOCTECDSKSIKSKLVLVHLRIHTGEKPYKCECDKSFQSOST 188
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 96 LRRHKILHTGEKPKFCHLCNVAACORDALTHSVKPYKCFGCRSYKQRSSLEEH 155
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 LVVHQRTHTGERPQCSEKSFYSYHFAVVVHRTHTGEKPYKCSMCDKAYSQRSNLKLH 248
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 156 KE-----RC-RTFLQSTDPGDTASAEARHIKAEMGSEKALVLDRIASNV 198
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

A:Reference number: A90892; MUID:87131089
A:Accession: B26595
A:Molecule type: DNA
A:Residues: 75-347 <CH2>
C:Cross-references: GB:M15709; NID:g193351; PID:g193352
C:Genetics:
A:Gene: mkr2
C:Keywords: DNA binding; zinc finger

Query Match 25.8%; Score 290; DB 2; Length 347;
Best Local Similarity 34.8%; Pred. No. 3.1e-18;
Matches 54; Conservative 34; Mismatches 61; Indels 6; Gaps 2;

QY 2 RDENVLSEPMGNAEPEIPYSYSREYENIKLERHVVSFDSSRPTSKMCMDCVGLS 61
DB 111 QSSLLKHQRIHTGKPYTCNVCDKIFERSLTV--HQRTHTGKPE---YKCECGKA 164
OY 62 CISFNVLMVHKRSHTGRPFQCNOCGASFTQKGNLLRHKLHTGKPEKCHLCNYACQRR 121
DB 165 FSQSMNLTVHQRHTGKPEKPYQCKEKGKAFKNSLSIQHERIHGKPYKCHDCGKAFSKN 224
QY 122 DALTGHLRHSVSKPYKCEFCGRSYKORSLSLEHK 156
DB 225 SSLTQHRRIHTGKPEKPYECMTCGRHFTGRSSLTVHQ 259

RESULT 11

zinc finger protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Aug-1997
C:Accession: I54338
R:Tommerup, N.; Aagaard, L.; Lund, C.L.; Boel, E.; Baxendale, S.; Bates, G.P.; Lehrach, Hum. Mol. Genet. 2, 1571-1575, 1993
A:Title: A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate gene for the W
A:Reference number: I54338; MUID:94093541
A:Accession: I54338
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-474 <RES>
A:Cross-references: GB:L15309; NID:g347905; PID:g347906
C:Genetics:
A:Gene: GDB:ZNF141
A:Cross-references: GDB:137046; OMIM:194648
A:Map position: 4p16.3-4p16.3

Query Match 25.7%; Score 289; DB 2; Length 474;
Best Local Similarity 34.6%; Pred. No. 5.2e-18;
Matches 56; Conservative 22; Mismatches 60; Indels 24; Gaps 2;

QY 16 EEPPEIPYSYSREYENI-----KLERHVVSFDSSRPTSKMCMC 55
DB 202 EECGKAFKWSLIPNEHKRIHTGKPYTCNOCGASFTQKGNLLRHKLHTGKPEKCHLCN 115
QY 56 DVCGLSCISFNVLMVHKRSHTGRPFQCNOCGASFTQKGNLLRHKLHTGKPEKCHLCN 115
DB 258 EECGKAFNRTTLTKHRIHAGEKPYTCNOCGASFTQKGNLLRHKLHTGKPEKCHLCN 317
QY 116 YACQRRDALTGHLRHSVSKPYKCEFCGRSYKORSLSLEHK 157
DB 318 KAFNRSTTLTKHRIHAGEKPYTCNOCGASFTQKGNLLRHKLHTGKPEKCHLCN 359

RESULT 12

finger protein OZF, Krueppel-related - human
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-1999
C:Accession: S43826
R:le Chalony, C.; Prosperi, M.T.; Haluza, R.; Apio, F.; Dutrillaux, B.; Goubin, G. J. Mol. Biol. 236, 399-404, 1994

A:Title: The OZF gene encodes a protein consisting essentially of zinc finger motifs.
A:Reference number: S43826; MUID:94149744
A:Accession: S43826
A:Molecule type: mRNA
A:Residues: 1-292 <LEG>
C:Cross-references: EMBL:X70394; NID:g468707; PID:g468708
C:Genetics:
A:Gene: GDB:ZNF146; OZF
A:Cross-references: GDB:230289; OMIM:601505
A:Map position: 19q13.1-19q13.1
C:Keywords: DNA binding; zinc finger

Query Match 25.6%; Score 288; DB 2; Length 292;
Best Local Similarity 30.4%; Pred. No. 3.9e-18;
Matches 62; Conservative 44; Mismatches 68; Indels 30; Gaps 5;

QY 4 ENVLKSEPMGNAEPEIPYSYSREYENIKLERHVVSFDSSRPTSKMCMDCVGLS 61
DB 85 ENLLTHQRIHTGKPEKPYQCKEKGKAFKNSLSIQHERIHGKPYKCHDCGKAFSKN 142
QY 55 -----CDVGLSCISFNVLMVHKRSHTGRPFQCNOCGASFTQKGNLLRH 99
DB 143 LTHEKEKIHTGKPEKPYQCKEKGKAFKNSLSIQHERIHGKPYKCHDCGKAFSKN 202
QY 100 IKLHTGKPEKCHLCNYACQRRDALTGHLRHSVSKPYKCEFCGRSYKORSLSLEHK 159
DB 203 VRIHSGDKPYECNCGKAFKNSLSIQHERIHGKPYKCHDCGKAFSKN 261
QY 160 RTF---LQSTDPGDTASAEARHIK 180
DB 262 HTGKPYQCECGKAFKNSLSIQHERIHGKPYKCHDCGKAFSKN 285

RESULT 13

finger protein ZNF11B - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S33989; S33994
R:Tunnacliffe, A.; Liu, L.; Moore, J.K.; Leversha, M.A.; Jackson, M.S.; Papi, L.; Fer Nucleic Acids Res. 21, 1409-1417, 1993
A:Title: Duplicated KOX zinc finger gene clusters flank the centromere of human chrom
A:Reference number: S30238; MUID:93219119
A:Accession: S33989
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 65-287 <TU2>
A:Cross-references: EMBL:X68685; NID:g296453; PID:g938236
A:Accession: S33994
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <TUN>
C:Cross-references: EMBL:X68684
C:Genetics:
A:Gene: GDB:ZNF11B
A:Cross-references: GDB:128306; OMIM:194522
A:Map position: 10q11.2-10q11.2
C:Keywords: zinc finger

Query Match 25.6%; Score 288; DB 2; Length 378;
Best Local Similarity 49.0%; Pred. No. 5.1e-18;
Matches 50; Conservative 20; Mismatches 32; Indels 0; Gaps 0;

QY 55 CDVGLSCISFNVLMVHKRSHTGRPFQCNOCGASFTQKGNLLRHKLHTGKPEKCHLC 114
DB 256 CNECGKAFWEKSHLTRHORVHTGKPEKCHLCNYACQRRDALTGHLRHSVSKPYKCEFCGRSYKORSLSLEHK 315
QY 115 NYACQRRDALTGHLRHSVSKPYKCEFCGRSYKORSLSLEHK 156
DB 316 GKAFSHKSLTLHQRTHTGKPEKPYQCNACGKTFYQKSDLTQKHQ 357

Db 75 NHCGVLSHKQGLDHTHTGKPYECNECGIAFSQKSHLVLRHTHTGKPYECQCG 134
QY 116 YACORRDLTGLTHSVKPYKCFGRSKYKQSSLEEH-----KERCRTFL 163
Db 135 KAHGKHALTDHLRIHTGKPYKNECGKTRHSSNLMQHLRSHTGKPYECKGKFSR 194
QY 164 QSTDPGDTASAEARHIKAEMG 184
Db 195 YN-----SSLTEHVTRHTG 208
RESULT 7
A:040751
finger protein MZF1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 24-Sep-1998
C:Accession: A40751
R:Hromas, R.; Collins, S.J.; Hickstein, D.; Raskind, W.; Deaven, L.L.; O'Hara, P.; Hagen
J. Biol. Chem. 266, 14183-14187, 1991
A:Title: A retinoic acid-responsive human zinc finger gene, MZF-1, preferentially expres
A:Reference number: A40751; MUID:913117761
A:Accession: A40751
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <HRO>
A:Cross-references: GB:M58297; NID:g189043; PID:g189044
C:Genetics:
A:Gene: GDB:ZNF42; MZF-1
A:Cross-references: GDB:125898; OMIM:194550
A:Map position: 19q13.2-19q13.4
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 26.4%; Score 297; DB 2; Length 485;
Best Local Similarity 41.0%; Pred. No. 1.1e-18;
Matches 57; Conservative 22; Mismatches 50; Indels 10; Gaps 2;

QY 18 PEIPYSYSREYENIKLERHVVSFDSRPTSGKMNCDVCGLSICISFNVLVHKKRSHTG 77
Db 351 PRCGGRFS-----QRLALTRHQRHTGKPY-----YHCGECGLGFTQVSRUTEHQRIHTG 400
QY 78 EPPFCNCGGASFTQKGNLRLHKLHTGKPKPKCHLCNYACQRRDALTGHLRTHSVKPY 137
Db 401 ERPFACPECGSGFRGHANLTHRRHTGKPYKPYKPYKPYKPYKPYKPYKPYKPYKPYK 460
QY 138 KCEFCGRSKYKQSSLEEH 156
Db 461 ACQDCGRFRPHOSTKLIQHQ 479

RESULT 8
S35305
zinc finger protein ZNF91 - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: S35305
R:Bellevue, E.J.; Marine, J.C.; Ried, T.; Lecocq, P.J.; Riviere, M.; Amemiya, C.; Pond
EMBO J. 12, 1363-1374, 1993
A:Title: Clustered organization of homologous KRAB zinc-finger genes with enhanced expre
A:Reference number: S35305; MUID:93223677
A:Accession: S35305
A:Molecule type: mRNA
A:Residues: 1-1191 <BEL>
A:Cross-references: EMBL:L11672; NID:g186773; PID:g186774
A:Note: the authors translated the codon GCA for residue 750 as Thr and GCT for residue
C:Genetics:
A:Gene: GDB:ZNF91; HPF7; HTF10
A:Cross-references: GDB:132284
A:Map position: 19p12-19p12
C:Keywords: DNA binding; zinc finger

Query Match 26.3%; Score 296; DB 2; Length 1191;

Best Local Similarity 34.9%; Pred. No. 3.3e-18;
Matches 66; Conservative 29; Mismatches 78; Indels 16; Gaps 4;
QY 12 MGNAEPEIPYSYSREYENIKLERHVVSFDSRPTSGKMNCDVCGLSICISFNVLVH 71
Db 1015 MHTGKPYKCEGCKAFN--RSSKLTHTKIHTGKPY-----YKCEGCKAFISSSTLNGH 1068
QY 72 KRSHTGERPFCNCGGASFTQKGNLRLHKLHTGKPKPKCHLCNYACQRRDALTGHLRTH 131
Db 1069 KRHTREKPYKCEGCKAFSSSTLTHRLHKLHTGKPKPKCHLCNYACQRRDALTGHLRTH 1128
QY 132 SVKPYKCFGRSKYKQSSLEEHKERCRTFLQSTDPGDTASAEARHIKAEMSERALVL 191
Db 1129 TGEKPYKCEKCKAFNQSILTNHKK-----IHTITVPIPLLWE-----AEAGSGRGOEM 1178
QY 192 DRLASNAVK 200
Db 1179 ETILANTVK 1187
RESULT 9
S22954
finger protein zfp-37 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C:Accession: S22954
R:Burke, P.S.; Wolgemuth, D.J.
Nucleic Acids Res. 20, 2827-2834, 1992
A:Title: zfp-37, a new murine zinc finger encoding gene, is expressed in a developmen
A:Reference number: S22954; MUID:92310982
A:Accession: S22954
A:Molecule type: mRNA
A:Residues: 1-553 <BUR>
A:Cross-references: EMBL:X64413; NID:g55474; PID:g55475
C:Superfamily: zinc finger protein zfp-36
C:Keywords: DNA binding

Query Match 26.1%; Score 294; DB 2; Length 553;
Best Local Similarity 38.0%; Pred. No. 2.2e-18;
Matches 57; Conservative 29; Mismatches 44; Indels 20; Gaps 3;

QY 48 PTKSK-MNCDVCGLSICISFNVLVHKKSHGTERPFCNCGGASFTQKGNLRLHKLHTGE 106
Db 208 PGSGKPYECNCHGKVLSHKQGLDHTHTGKPYKPYKPYKPYKPYKPYKPYKPYKPYK 267
QY 107 KPEKCHLCNYACQRRDALTGHLRTHSVKPYKCFGRSKYKQSSLEEH----- 156
Db 268 KPYECGCKAHGKHKHALTDHLRIHTGKPYKNECGKTRHSSNLMQHLRSHTGKPYE 327
QY 156 -KERCRTFLQSTDPGDTASAEARHIKAEMG 184
Db 328 CKEGCKSFYN-----SSLTEHVTRHTG 350

RESULT 10
S00549
developmental control protein Kr2 - mouse
N:Alternate names: mkr2
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 20-Mar-1998
C:Accession: S00549; B26595
R:Chowdhury, K.; Dressler, G.; Breier, G.; Deutsch, U.; Gruss, P.
EMBO J. 7, 1345-1353, 1988
A:Title: The primary structure of the murine multifinger gene mkr2 and its specific e
A:Reference number: S00549; MUID:88312581
A:Accession: S00549
A:Molecule type: mRNA
A:Residues: 1-347 <CHO>
A:Cross-references: EMBL:Y00850; NID:g52810; PID:g52811
R:Chowdhury, K.; Deutsch, U.; Gruss, P.
Cell 48, 771-778, 1987
A:Title: A multigene family encoding several "finger" structures is present and diffe

Db 114 HKYCGRSYKQSSLEEKHERCHNYLESGLPGVCPVKEETNHNHMAEDLCKIGAERS 173
 QY 189 LVLDRLASNVAKRKSSMPQKF 209
 Db 174 LVLDRLASNVAKRKSSMPQKF 194

RESULT 3
 B56229
 lymphoid transcription factor Ikaros/Lyf-1, form V - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
 C:Accession: B56229
 R:Haum, K.; Ernst, P.; Lo, K.; Kim, G.S.; Turk, C.; Smale, S.T.
 Mol. Cell. Biol. 14, 7111-7123, 1994
 A:Title: The lymphoid transcription factor Lyf-1 is encoded by specific, alternatively spliced transcripts
 A:Reference number: A56229; MUID:95021239
 A:Accession: B56229
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-427 <HA>
 A:Cross-references: GB:S74518
 C:Keywords: alternative splicing

Query Match 50.1%; Score 564; DB 2; Length 427;
 Best Local Similarity 75.9%; Pred. No. 4.6e-42;
 Matches 107; Conservative 11; Mismatches 15; Indels 8; Gaps 2;

QY 77 GERPFQNCOGASFTQKGNLLRHKLHTGEKPKFKCHLCNYACORRDLTGLHRTSHVSRP 136
 Db 54 GERPFQNCOGASFTQKGNLLRHKLHTGEKPKFKCHLCNYACORRDLTGLHRTSHVSRP 113

QY 137 YKCEFGRSYKORSSLEEKHERCRFTLQSTD-PG-----DTASAEARHIKAEMGSERA 188
 Db 114 HKYCGRSYKORSSLEEKHERCHNYLESGLPGVCPVKEETNHNHMAEDLCKIGAERS 173

QY 189 LVLDRLASNVAKRKSSMPQKF 209
 Db 174 LVLDRLASNVAKRKSSMPQKF 194

RESULT 4
 S00647
 finger protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
 C:Accession: S00647
 R:Ruiz i Altaba, A.; Perry-O'Keefe, H.; Melton, D.A.
 EMBO J. 6, 3065-3070, 1987
 A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
 A:Reference number: S00647; MUID:88082679
 A:Accession: S00647
 A:Molecule type: mRNA
 A:Residues: 1-1350 <RUI>
 A:Cross-references: EMBL:X06021
 A:Note: it is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the initiator
 C:Genetics: fin
 C:Keywords: DNA binding; zinc finger

Query Match 28.2%; Score 317.5; DB 2; Length 1350;
 Best Local Similarity 36.9%; Pred. No. 4.9e-20;
 Matches 66; Conservative 33; Mismatches 59; Indels 21; Gaps 4;

QY 4 ENVLKSEPMGNAEPEIPYSYREYNEYNKLERHVVSFDSSRPTSGKMN---CDVGL 60
 Db 308 DSVGTDPDLSQNVASSPYSCSKRKTFRWK-----SFLNHOOTHREKPYLCSHCKN 361

QY 61 SCISFNVLVHVRSHGTGERPFCNOCGASFTQKGNLLRHKLHTGEKPKFKCHLCNYACOR 120
 Db 174 LVLDRLASNVAKRKSSMPQKF 194

Db 362 GFIONSDLVKHFRTHTGERPYQCAECHKGFIOKSDLVKHLRHTHTGEKPKFKSCHCKKFFE 421
 QY 121 RDALTGHLRTHSVKPKCEFCGRSYKORSSLEEKHE-----RC-----RTFLQSTD 167
 Db 422 RSALAKHQHTHTGERPKYKSCDCGKEFTQRSNLLHQRIHTGERPKYKCTLCDRTFFIONSD 480

RESULT 5
 S26823
 finger protein ZNF43 - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1998
 C:Accession: S26823
 R:Lowering, R.; Trowsdale, J.
 Nucleic Acids Res. 19, 2921-2928, 1991
 A:Title: A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell
 A:Reference number: S26823; MUID:91279444
 A:Accession: S26823
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-803 <LOW>
 A:Cross-references: EMBL:X59244; NID:g38031; PID:g38032
 C:Genetics:
 A:Gene: GDB:ZNF43; HTF6
 A:Cross-references: GDB:128653
 A:Map position: 19p13.1-19p12
 C:Keywords: DNA binding

Query Match 26.6%; Score 299; DB 2; Length 803;
 Best Local Similarity 35.3%; Pred. No. 1.2e-18;
 Matches 61; Conservative 26; Mismatches 56; Indels 30; Gaps 3;

QY 4 ENVLKSEPMGNAEPEIPYSYREYNEYNKLERHVVSF 43
 Db 473 EKPYKCECGKA-----FSRSSNLTKKHKTIEKKPKYKCECGKAFKWSKLTETKITH 526

QY 44 DSSRPTSGKMNCDVGLSCISFNVLVHVRSHGTGERPFCNOCGASFTQKGNLLRHKLH 103
 Db 527 TGEKP-----YKCECGKAFNHFSLTKHRIHTGEKPKYKCECGKAFKWSKLTETKITH 582

QY 104 TGEKPKFKCHLCNYACORRDLTGLHRTSHVSRPFCGRSYKORSSLEEK 156
 Db 583 TGEKPKCECGKAFKWSKLTETKITHTKHRIHTGEKPKYKCECGKAFKWSKLTETKITH 635

RESULT 6
 S10245
 finger protein, testis - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 10-Sep-1997
 C:Accession: S10245
 R:Neik, D.; Dudley, K.; Cunningham, P.; Akhavan, M.
 Nucleic Acids Res. 18, 3655, 1990
 A:Title: Cloning and sequencing of a zinc finger cDNA expressed in mouse testis.
 A:Reference number: S10245; MUID:90301500
 A:Accession: S10245
 A:Molecule type: mRNA
 A:Residues: 1-411 <NEL>
 A:Cross-references: EMBL:X52533; NID:g53456; PID:g53457
 C:Keywords: DNA binding; zinc finger

Query Match 26.5%; Score 298.5; DB 2; Length 411;
 Best Local Similarity 32.8%; Pred. No. 6.7e-19;
 Matches 66; Conservative 34; Mismatches 66; Indels 35; Gaps 5;

QY 6 VLKSEPMGNAEPEIPYSYREYNEYNKLERHVVSFDSSRPTSGK-MNC 55
 Db 21 VTKTKPAKNVRK-----YARHSASHTKEDKIQTGEKRSKCHRTSPKPEKAGSGKPYEC 74

QY 56 DVCGLSCISFNVLVHVRSHGTGERPFCNOCGASFTQKGNLLRHKLHTGEKPKFKCHLCN 115
 Db 174 LVLDRLASNVAKRKSSMPQKF 194

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:08:51 ; Search time 18.18 Seconds
(without alignments)
460.599 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 1126
Sequence: 1 ERDENYKSEPMGNAEPEI.....VLDRLASNVAKRKSSMPQKF 209

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	671	59.6	518	2 A56355	DNA-binding protei
2	564	50.1	431	2 I59572	Ikaros DNA binding
3	564	50.1	427	2 B56329	lymphoid transcrip
4	317.5	28.2	1350	2 S00647	finger protein - A
5	299	26.6	803	2 S26823	finger protein ZNF
6	298.5	26.5	411	2 S10325	finger protein, te
7	297	26.4	485	2 A40751	finger protein MZF
8	296	26.3	1191	2 S35305	zinc finger protei
9	294	26.1	553	2 S22954	finger protein zfp
10	290	25.8	347	2 S00549	developmental cont
11	289	25.7	474	2 I54338	zinc finger protei
12	288	25.6	292	2 S43826	finger protein O2F
13	288	25.6	378	2 S33894	finger protein ZNF
14	288	25.6	595	2 G02075	transcription repr
15	287.5	25.5	439	2 S06556	finger protein (cl
16	285.5	25.4	614	2 JH0500	zinc finger protei
17	284	25.2	732	2 S47073	finger protein H2F
18	282.5	25.1	510	2 A36901	chorion transcript
19	282.5	25.1	428	2 A32891	finger protein, l
20	282.5	25.1	580	2 A37107	spermatogenesis pr
21	282	25.0	336	2 S06578	finger protein (cl
22	282	25.0	393	2 JN0533	finger protein pML
23	281.5	25.0	427	2 A35659	krueppel-related p
24	281.5	25.0	469	2 I38600	zinc finger protei
25	280.5	24.9	589	2 I38598	zinc finger protei
26	280	24.9	543	2 B34612	zinc finger protei
27	279	24.8	247	2 S06553	finger protein (cl
28	279	24.8	576	2 A48157	renal transcriptio
29	278.5	24.7	229	2 A48927	Kruppel-like zinc
30	278.5	24.7	428	2 S03677	finger protein (cl
31	278	24.7	247	2 S47070	finger protein H2F
32	278	24.7	636	2 I48689	gene NK10 protein
33	277	24.6	519	2 A38073	transcription acti
34	277	24.6	693	2 I37570	zinc finger protei
35	277	24.6	295	2 A46017	zinc finger protei
36	276.5	24.6	546	2 I49636	DNA-binding protei
37	276.5	24.6	242	2 S30238	finger protein ZNF
38	275.5	24.5	491	2 S00733	finger protein HF
39	275	24.4	453	2 S06544	finger protein (cl

40 275 24.4 591 2 S65088 finger protein XFO
41 274 24.3 686 2 A34612 zinc finger protei
42 273.5 24.3 169 2 A39240 finger protein mfg
43 273 24.2 651 2 B32891 finger protein 2,
44 272.5 24.2 210 2 S06577 finger protein (cl
45 271.5 24.1 462 2 S10397 finger protein kox

ALIGNMENTS

RESULT 1
A56355
DNA-binding protein Ikaros form 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C:Accession: A56355
R: Molnar, A.; Georgopoulos, K.
Mol. Cell. Biol. 14, 8292-8303, 1994
A:Title: The Ikaros gene encodes a family of functionally diverse zinc finger DNA-bin
A:Reference number: A56355; MUID:95059058
A:Accession: A56355
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-518 <MOL>
C:Keywords: alternative splicing; DNA binding

Query Match 59.6%; Score 671; DB 2; Length 518;
Best Local Similarity 73.7%; Pred. No. 2.6e-51;
Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

QY 47 RPTSGKMGNDVCGLSGICISFNVLVHVRSHSGTGERPFCNCGASFTQKGNLRLHKLHTGE 106
DB 111 RLPNGKLKCDICGIVCGPNVLMVHVRSHSGTGERPFCNCGASFTQKGNLRLHKLHSGE 170
QY 107 KPFCHLCNVACORRDALTGHLRTHSVKPYKCFGCGRSYKQSSLEEHKERCRTFLOST 166
DB 171 KPFCHLCNVACRRDALTGHLRTHSVGKPKGCGYGRSYKQSSLEEHKERCERNYLESM 230
QY 167 D-PG-----DTASAEARHKAEMGERALVLDRLASNVAKRKSSMPQKF 209
DB 231 GLGPMYPIVKEETNHNHNEAEDLCKIGAESRLVLDRLASNVAKRKSSMPQKF 281

RESULT 2
Ikaros DNA binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C:Accession: I59572
R: Georgopoulos, K.; Moore, D.D.; Derfler, B.
Science 258, 808-812, 1992
A:Title: Ikaros, an early lymphoid-specific transcription factor and a putative media
A:Reference number: I59572; MUID:93068267
A:Accession: I59572
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-431 <RES>
A:Cross-references: GB:I03547; NID:g198286; PID:g198287
C:Genetics:
A:Gene: Ikaros

Query Match 50.1%; Score 564; DB 2; Length 431;
Best Local Similarity 75.9%; Pred. No. 4.7e-42;
Matches 107; Conservative 11; Mismatches 15; Indels 8; Gaps 2;

QY 77 GERPFQNCQCGASFTQKGNLRLHKLHTGKPKFCHLCNVACORRDALTGHLRTHSVKEKP 136
DB 54 GERPFQNCQCGASFTQKGNLRLHKLHSGEKPKFCHLCNVACRRDALTGHLRTHSVGKP 113
QY 137 YKCFGCGRSYKQSSLEEHKERCRTFLOSTD-PG-----DTASAEARHKAEMGSERA 188

MOLECULE TYPE: protein
US-08-398-590A-40

Query Match
Best Local Similarity 20.0%; Score 225; DB 2; Length 676;
Matches 46; Conservative 32; Mismatches 73; Indels 14; Gaps 2;

OY 1 ERDENLVKSEPMGNAPEEPIYS-----YGREYNEYENIKLERHVVSFDSRPTS 50
Db 270 ESAEQAKARESGSTAEGDFSGPIRCDRCGYNTNRVDYHTAHLKHHTHTRAGDNER--- 327

OY 51 GWNCDVCLSCISFNVLVHVRSHRTGERPFCNOCGASFTQGNLLRHILKHTGKPEK 110
Db 327 -VYKCIITTYTVSEYHVKHRLNHFPRKYVTCGKCNFYSDRKNYVQHVRTHTGERPYK 385

OY 111 CHLCNYACORRDALGHLRTHSVKPYKCFGRSYKORSSLEEH 155
Db 386 CELCPYSSOKTHLRHMRTHSGERPFKDCQCSYVASNOHEVTRH 430

RESULT 14
US-08-718-661-2
Sequence 2, Application US/08718661
Patent No. 5876972
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,661
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 667 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-661-2

Query Match
Best Local Similarity 19.9%; Score 224; DB 2; Length 667;
Matches 50; Conservative 25; Mismatches 70; Indels 12; Gaps 5;

OY 55 CDVCGLSICISFNVLVHVRSHRTGERPFCN--OCGASFTQGNLLRHILKHTGKPEKCH 112
Db 6 CQCKGRSFVTEKFTIHNYSHSRERPFKCSKAECKGAFVSKYKLMRHMATHSPQKHQCT 65

OY 113 LCNYACORRDALGHLRTHSVK-PYKCFGRSYKORSSLEEHKERCRTFLQSTDPGDT 171
Db 66 HCEKTFNRKDLKLNHLQTHDPNKSISYACDDCGKKYHMLGYKRH-----LALHSASGDL 120

OY 172 ASAEARHIKAEMGSEALVLDRLASNAVKRKSMPQK 208
Db 121 TCGVC---TLELGSFVL-LDHLKSHAEEKANQAPRE 153

RESULT 15
US-08-224-482-4
Sequence 4, Application US/08224482
Patent No. 5837692
GENERAL INFORMATION:
APPLICANT: Mercola, Dan
APPLICANT: Adamson, Eileen D.
TITLE OF INVENTION: Inhibition of the Mitogenic Activity of

TITLE OF INVENTION: PDGF by Mammalian EGR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/224,482
FILING DATE: 07-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-ME 9913
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-224-482-4

Query Match
Best Local Similarity 18.7%; Score 210.5; DB 2; Length 543;
Matches 40; Conservative 19; Mismatches 46; Indels 5; Gaps 2;

OY 78 ERPFQC--NOCGASFTQGNLLRHILKHTGKPEKCHLCNYACORRDALGHLRTHSVK 135
Db 335 ERPYACPVESCDRRFSRSDLTIRIHTGQKPFQCRICMNFESRSDLTHTHTGTCEK 394

OY 136 PYKCFGRSYKORSSLEEHKERCRTFLQSTDPGDTASAEARHIKAEMGS 185
Db 395 PFACDICGRKFARS---DERKRHTKIHLRQKDKKADKSVASSATSSLS 441

Search completed: November 6, 1999, 09:41:36
Job time: 68 sec

```

RESULT 13
US-08-398-590A-40
; Sequence 40, Application US/08398590A
; Patent No. 5935811
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Schoenherr, Christopher J.
; TITLE OF INVENTION: Neuron-Restrictive Silencer Factor
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,590A
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,445
; FILING DATE: 06-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60897/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 676 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```


TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-170A-10

Query Match 23.9%; Score 269; DB 2; Length 711;
Best Local Similarity 40.7%; Pred. No. 2e-22;
Matches 48; Conservative 22; Mismatches 48; Indels 0; Gaps 0;
QY 39 HVVSDSSRPTSGKMNCDVGLSCISFNVLVHKRSHGTGERPQCNOQGASFTQKGNLLR 98
DB 461 HLIVHQRSHGTGKPYOCHNCGKSFISKSLDIHRIHTGKPYECSDCGTKFTQKSHLNI 520
QY 99 HIKLHTGEKPFKCHLCNYACQRDALTGHLRTHSVKPKYKCFGCRSYKORSLEEHK 156
DB 521 HOKIHTGERHVGSECGKAFNQKSIILSMHQRIHTGKPYKCEGKRAFTSKSOFKEHQ 578

RESULT 7
US-08-933-750C-17
Sequence 17, Application US/08933750C
Patent No. 5932442
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRSTNOT03
CLONE: 641127
US-08-933-750C-17

Query Match 23.4%; Score 263; DB 2; Length 488;
Best Local Similarity 28.1%; Pred. No. 5.6e-22;
Matches 62; Conservative 41; Mismatches 72; Indels 46; Gaps 7;
QY 6 VLKSPMGNAEPEIPYSYSREY-----NEYENIKLERHVVSFDSRPTSGKMNCDVC 58
DB 187 ILNTSPDGHPIEK--IHTAKQYEGSCQCKNFSSQSELLLHORDHTEKP-----YKCEOC 240
QY 59 GLSCISFNVLVHKRSHGTGERPQCNOQGASFTQKGNLLRHLRIKHTGKPKFCHLCNYAC 118
DB 241 GKGFTRSSLLIHAQVHTDEKPYKDKCGKGFTRSSLLIHHAVHTGKPKYKCDKCGKGF 300
QY 119 QRDALTGHLRTHSVKPKYKCFGCRSYKORSLEEHK-----RC-----RTFLQST 166
DB 301 SOSKSLHRIHQRVHTGKPYECECGMSFSQRSNLHIHQRVHTGTERPYKCGCGKGFSSQSS 360
QY 167 D-----PGDTASAEAR-HIKAEMGSE 186
DB 361 NLHRIHRCIHTGKPYQCYECGKGFSSQSSDLRIHLRVHTGK 401

RESULT 8
US-08-074-967-2
Sequence 2, Application US/08074967
Patent No. 5641672
GENERAL INFORMATION:
APPLICANT: Dalia-Favera, Riccardo
APPLICANT: Chaganti, R.S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/074,967
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 706 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-074-967-2

Query Match 22.9%; Score 258; DB 1; Length 706;
Best Local Similarity 38.7%; Pred. No. 3.6e-21;
Matches 48; Conservative 18; Mismatches 54; Indels 4; Gaps 1;

```

1  RESULT          5
2  US-08-933-750C-14
3  Sequence 14, Application US/08933750C
4  Patent No. 5932442
5  GENERAL INFORMATION:
6  APPLICANT: Lal, Preeti
7  APPLICANT: Hillman, Jennifer L.
8  APPLICANT: Bandman, Olga
9  APPLICANT: Shah, Purvi
10 APPLICANT: Au-Young, Janice
11 APPLICANT: Yue, Henry
12 APPLICANT: Guegler, Karl J.
13 APPLICANT: Corley, Neil C.
14 TITLE OF INVENTION: HUMAN REGULATORY
15 NUMBER OF SEQUENCES: 98
16 CORRESPONDENCE ADDRESS:
17 ADDRESSEE: Incyte Pharmaceutical
18 STREET: 3174 Porter Drive
19 CITY: Palo Alto
20 STATE: CA
21 COUNTRY: USA
22 ZIP: 94304
23 COMPUTER READABLE FORM:
24 MEDIUM TYPE: Diskette
25 COMPUTER: IBM Compatible
26 OPERATING SYSTEM: DOS

```

```

1  RESULT 6
2  US-08-820-170A-10
3  ; Sequence 10, Application US/08820170A
4  ; Patent No. 5831058
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Tsutomu, FUJIWARA
7  ; APPLICANT: Takeshi, WATANABE
8  ; APPLICANT: Masato, HORIE
9  ; APPLICANT: Toyomasa, KATAGIRI
10 ; TITLE OF INVENTION: HUMAN GENE
11 ; NUMBER OF SEQUENCES: 42
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: Sughrue, Mion, Zinn,
14 ; STREET: 2100 Pennsylvania Avenue
15 ; CITY: Washington
16 ; STATE: D.C.
17 ; COUNTRY: United States
18 ; ZIP: 20037-3202
19 ; COMPUTER READABLE FORM:
20 ; MEDIUM TYPE: Floppy disk
21 ; COMPUTER: IBM PC compatible
22 ; OPERATING SYSTEM: PC-DOS/MS-DOS
23 ; SOFTWARE: PatentIn Release #1.0
24 ; CURRENT APPLICATION DATA:
25 ; APPLICATION NUMBER: US/08/820.1
26 ; FILING DATE:
27 ; CLASSIFICATION: 536
28 ; TELECOMMUNICATION INFORMATION:

```



```

RESULT      4
US-08-933-750C-4
; Sequence 4, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

Result No.	Score	Query Match	Length	DB	ID	Description	
1	679.5	60.3	470	2	US-08-465-590-153	Sequence 153, Appl	
2	679.5	60.3	470	3	PCI-US95-09345-13	Sequence 13, Appl	
3	679	60.3	568	3	PCI-US93-08743-5	Sequence 5, Appl	
4	273.5	24.3	338	2	US-08-9333-750C-4	Sequence 4, Appl	
5	272	24.2	368	2	US-08-9333-750C-14	Sequence 14, Appl	
6	269	23.9	711	2	US-08-820-170A-10	Sequence 10, Appl	
7	263	23.4	488	2	US-08-9333-750C-17	Sequence 17, Appl	
8	258	22.9	706	1	US-08-074-967-2	Sequence 2, Appl	
9	258	22.9	706	2	US-08-553-541B-2	Sequence 2, Appl	
10	258	22.9	706	3	PCI-US94-06669-2	Sequence 2, Appl	
11	252.5	22.4	728	3	PCI-US95-08429-5	Sequence 5, Appl	
12	252.5	22.4	727	3	PCI-US95-08429-9	Sequence 9, Appl	
13	225	20.0	676	2	US-08-398-590A-40	Sequence 40, Appl	
14	224	19.9	667	2	US-08-718-661-2	Sequence 2, Appl	
15	210.5	18.7	543	2	US-08-224-483-4	Sequence 4, Appl	
16	210	18.7	387	2	US-08-224-482-8	Sequence 8, Appl	
17	209.5	18.6	533	2	US-08-040-548-1	Sequence 1, Appl	
18	209.5	18.6	533	2	US-08-466-344-1	Sequence 1, Appl	
19	209.5	18.6	486	2	US-08-224-482-2	Sequence 2, Appl	
20	208	18.5	84	2	US-08-459-568-62	Sequence 62, Appl	
21	208	18.5	84	2	US-08-399-411-62	Sequence 62, Appl	
22	207.5	18.4	456	2	US-08-040-548-2	Sequence 2, Appl	
23	207.5	18.4	456	2	US-08-466-344-2	Sequence 2, Appl	
24	207.5	18.4	455	2	US-08-224-482-6	Sequence 6, Appl	
25	206	18.3	89	2	US-08-040-548-8	Sequence 8, Appl	
26	206	18.3	89	2	US-08-466-344-8	Sequence 8, Appl	
27	204.5	18.2	429	1	US-08-234-783-4	Sequence 4, Appl	
28	204.5	18.2	429	1	US-08-456-907-4	Sequence 4, Appl	
29	204.5	18.2	429	3	PCI-US95-05523-4	Sequence 4, Appl	
30	204	18.1	547	2	US-08-340-203A-3	Sequence 3, Appl	
31	204	18.1	547	2	US-08-452-567-3	Sequence 3, Appl	
32	204	18.1	547	2	US-08-452-567-3	Sequence 3, Appl	
33	201.5	17.9	345	1	US-08-102-942A-2	Sequence 2, Appl	
34	199	17.7	675	1	US-08-317-522A-9	Sequence 9, Appl	
35	199	17.7	675	1	US-08-439-818A-9	Sequence 9, Appl	
36	199	17.7	675	2	US-08-751-963-9	Sequence 9, Appl	
37	199	17.7	675	2	US-08-738-975-9	Sequence 9, Appl	
38	199	17.7	675	2	US-08-728-626-9	Sequence 9, Appl	
39	194	17.2	429	1	US-08-102-942A-6	Sequence 6, Appl	

KW Alzheimer's disease; therapy; diagnosis.
OS Mus sp.

FT Key Location/Qualifiers
FT Region 119..139
FT FT /note= "zinc finger motif"
FT Region 147..167
FT FT /note= "zinc finger motif"
FT Region 175..195
FT FT /note= "zinc finger motif"
FT Region 203..224
FT FT /note= "zinc finger motif"
FT Region 460..480
FT FT /note= "zinc finger motif"
FT Region 491..513
FT FT /note= "zinc finger motif"

CA2194256-A.

05-MAR-1998.

02-JAN-1997; 194256.

PR 05-SEP-1996; US-711417.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K;

DR WPI: 98-378292/33.

DR N-PSDB: V42808.

PT New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and related products, used to treat e.g. immune diseases or cancer and to control cell differentiation

PS Claim 7: Page 75-77; 158pp; English.

CC This is the amino acid sequence of murine Ikaros protein isoform mtk-1, deduced from a cDNA clone (see V42808) obtained from a mature murine T cell line E14 library. Native Ikaros is active in the early stages of lymphocyte differentiation, binding to and activating the CD3-delta gene enhancer (see V42804). Proteins of the murine Ikaros family (see also W70963 and W70965-68) are isoforms that arise from differential splicing of Ikaros gene transcripts, and contain different combinations of zinc fingers.

CC They are expressed primarily in T cells in the adult and may play a role as a genetic switch regulating entry into the T cell lineage. The murine and human sequences (see W70964, W70969 and W70971) are very similar. The invention provides Ikaros nucleic acids, vectors and host cells expressing Ikaros polypeptides. These can be used to treat T and B cell diseases (e.g. immune deficiencies caused by drugs, radiation or cancer), to control expression of heterologous genes placed under control of an Ikaros-responsive element, to treat nervous system diseases (e.g. Alzheimer's disease) and to modulate cell division, amplification or differentiation, especially in haematopoietic cells. Some Ikaros isoforms are antagonistic of others and may be used to inhibit interaction with DNA sequences.

SQ Sequence 518 AA;

Query Match 59.4%; Score 669; DB 1; Length 518;

Best Local Similarity 73.7%; Pred. No. 4.6e-59;

Matches 126; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

Qy 47 RPTSGKMNCDVGLSCISFNVLWVHKRSHGTPFQCNQCGASFTQGNLLRHKLHTGE 106

Db 111 RLPNGKLCDCIGVIGPNVLWVHKRSHGTPFQCNQCGASFTQGNLLRHKLHSGE 170

Qy 107 KPFKCHLCNYACORRDALTGLHRLTHSVKPKYKCFECGRSYKQRSLEEHKRCRTFLOST 166

Db 171 KPFKCHLCNYACRRDALTGLHRLTHSVGPKHKGCGYKQRSLEEHKRCRCHNYLES 230

Qy 167 D-PG-----DTASAEARHIKAEMGSRALVLDRLASNVAKRKSSMPQKF 209

Db 231 GLPGVCPVKEETNHNMAEDLCKIGAERSLVLDRLASNVAKRKSSMPQKF 281

RESULT 15

R92014

ID R92014 standard; Protein; 431 AA.

AC R92014;

DT 08-MAY-1996 (first entry)

DE Murine Ikaros protein mtk-2.
KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte; immunocomprised; immune system disorder; nervous system disorder;
KW animal model; mtk-2.
OS Mus musculus.
PN WO9604372-A1.
PD 15-FEB-1996.
PR 28-JUL-1994; U09345.
PR 29-JUL-1994; US-283300.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI: 96-129389/13.
DR N-PSDB: T16059.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
PT Disclosure; Fig 1; 102pp; English.
CC The sequence of 48 kba mouse Ikaros protein mtk-2 (R92014) was deduced from mouse Ikaros cDNA (T16059) isolated from a mature T-cell line E15 library. Ikaros protein is a master regulator of hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92016-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.

SQ Sequence 431 AA;

Query Match 49.2%; Score 554; DB 1; Length 431;

Best Local Similarity 75.2%; Pred. No. 1.2e-47;

Matches 106; Conservative 11; Mismatches 16; Indels 8; Gaps 2;

Qy 77 GERPFQCNQCGASFTQGNLLRHKLHTGKPKCHLCNYACORRDALTGLHRLTHSVKPK 136

Db 54 GERPFQCNQCGASFTQGNLLRHKLHTGKPKCHLCNYACRRDALTGLHRLTHSVKPK 113

Qy 137 YKCEFCGRSYKQRSLEEHKRCRTFLOSTD-PG-----DTASAEARHIKAEMGSE 188

Db 114 HKGCGCRSYKQRSLEEHKRCRCHNYLESGLPGVCPVKEETNHNMAEDLCKIGAERS 173

Qy 189 LVLDRLASNVAKRKSSMPQKF 209

Db 174 LVLDRLASNVAKRKSSMPQKF 194

Search completed: November 6, 1999, 06:55:01

Job time: 83 sec


```
FT      /note= "unidentified amino acid"
FT misc_difference 430..432
FT      /note= "unidentified amino acids"
FT misc_difference 467
FT      /note= "unidentified amino acid"
FT misc_difference 469
FT      /note= "unidentified amino acid"
PN WO9604372-A1.
PD 15-FEB-1996.
PF 28-JUL-1995; U09345.
PR 29-JUL-1994; US-283300.
PA (GEO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI: 96-129389/13.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
PT severely immuno-compromised and can be used as model to determine
PT effects of treatment for immune and nervous system disorders
PS Disclosure; Page 75-76; 102pp; English.
CC The sequence of an Ikaros protein (R92021) is provided in the
CC specification. Ikaros protein is a master regulator of
CC hematopoietic differentiation and a major determinant in lymphocyte
CC differentiation. Isoforms of Ikaros (see R92014-19) arise from
CC differential splicing of Ikaros gene transcripts. Transgenic animals,
CC pref. mice, having a mutated Ikaros transgene, esp. a mutation that
CC alters the DNA binding domain of the Ikaros protein, are used as
CC models to determine the effects of treatments for immune or nervous
CC system disorders.
SQ Sequence 470 AA;

Query Match          60.3%; Score 679.5; DB 1; Length 470;
Best Local Similarity 73.3%; Pred. No. 3.6e-60;
Matches 126; Conservative 14; Mismatches 23; Indels 9; Gaps 2;

QY 47 RPTSGKMNDCVGLSCISFNVLVHKSHSGTGERPFCNQCGASFTQGNLRLHKLHTGE 106
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 RLPNGKLCDCIGXCIGPNVLVHKSHSGTGERPFCNQCGASFTQGNLRLHKLHSGE 119
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 107 KPFKCHLCNACRRDALTGLHRTSHVEKPKCEFGCRSYKQKSSLEHKERCRTFLQST 166
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 120 KPFKCHLCNACRRDALTGLHRTSHVGPKHKCGYGRSYKQKSSLEHKERCRTFLQST 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 167 D-PG-----DFASAEARHIKEMSERALVLDRLASNVAKRKSSMPQKF 209
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 180 GLPGXXXVPIKEETXHXEMAEDLCKIGXERSLVLDRLASNVAKRKSSMPQKF 231
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 9
W72678
ID W72678 standard; Protein; 470 AA.
AC W72678;
DT 14-JAN-1999 (first entry)
DE Ikaros protein general formula.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
KW Alzheimer's disease.
US Homo sapiens.
FH Key
FH Location/Qualifiers
FT Misc_difference 1 /note= "any amino acid"
FT Misc_difference 2 /note= "any amino acid"
FT Misc_difference 74 /note= "any amino acid"
FT Misc_difference 74 /note= "any amino acid"
FT Misc_difference 163 /note= "any amino acid"
FT Misc_difference 163 /note= "any amino acid"
FT Misc_difference 184 /note= "any amino acid"
FT Misc_difference 184 /note= "any amino acid"
FT Misc_difference 185 /note= "any amino acid"
FT Misc_difference 185 /note= "any amino acid"
FT Misc_difference 186 /note= "any amino acid"
FT Misc_difference 186 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"

FT      /note= "any amino acid"
FT Misc_difference 196
FT      /note= "any amino acid"
FT Misc_difference 207
FT      /note= "any amino acid"
FT Misc_difference 236
FT      /note= "any amino acid"
FT Misc_difference 240
FT      /note= "any amino acid"
FT Misc_difference 246
FT      /note= "any amino acid"
FT Misc_difference 251
FT      /note= "any amino acid"
FT Misc_difference 252
FT      /note= "any amino acid"
FT Misc_difference 255
FT      /note= "any amino acid"
FT Misc_difference 261
FT      /note= "any amino acid"
FT Misc_difference 285
FT      /note= "any amino acid"
FT Misc_difference 300
FT      /note= "any amino acid"
FT Misc_difference 301
FT      /note= "any amino acid"
FT Misc_difference 302
FT      /note= "any amino acid"
FT Misc_difference 304
FT      /note= "any amino acid"
FT Misc_difference 306
FT      /note= "any amino acid"
FT Misc_difference 316
FT      /note= "any amino acid"
FT Misc_difference 319
FT      /note= "any amino acid"
FT Misc_difference 320
FT      /note= "any amino acid"
FT Misc_difference 329
FT      /note= "any amino acid"
FT Misc_difference 331
FT      /note= "any amino acid"
FT Misc_difference 352
FT      /note= "any amino acid"
FT Misc_difference 367
FT      /note= "any amino acid"
FT Misc_difference 368
FT      /note= "any amino acid"
FT Misc_difference 369
FT      /note= "any amino acid"
FT Misc_difference 371
FT      /note= "any amino acid"
FT Misc_difference 372
FT      /note= "any amino acid"
FT Misc_difference 373
FT      /note= "any amino acid"
FT Misc_difference 374
FT      /note= "any amino acid"
FT Misc_difference 375
FT      /note= "any amino acid"
FT Misc_difference 380
FT      /note= "any amino acid"
FT Misc_difference 384
FT      /note= "any amino acid"
FT Misc_difference 385
FT      /note= "any amino acid"
FT Misc_difference 397
FT      /note= "any amino acid"
FT Misc_difference 407
FT      /note= "any amino acid"
FT Misc_difference 430
FT      /note= "any amino acid"
FT Misc_difference 431
FT      /note= "any amino acid"
```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 06:55:00 ; Search time 21.23 Seconds
(without alignments)
233.179 Million cell updates/sec

Title: US-09-019-348-8
Perfect score: 126
Sequence: 1 ERDENVLKSEPMGNAEPEI.....VLDRLASNVAKRKSMPQKF 209

Scoring table: BLOSOM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1126	100.0	209	1 W15575	Human Aiolos polyp
2	1020.5	90.6	507	1 W15574	Mouse Aiolos polyp
3	682.5	60.6	537	1 R46964	Peptide with Ikaros
4	682.5	60.6	461	1 R20215	Human Ikaros prote
5	682.5	60.6	461	1 W2672	Human Ikaros. Ikar
6	682.5	60.6	461	1 W70964	Human Ikaros isofo
7	682.5	60.6	516	1 W0971	Human Ikaros isofo
8	679.5	60.3	470	1 R2021	Ikaros protein. Tr
9	679.5	60.3	470	1 W2678	Ikaros protein gen
10	679	60.3	568	1 R46965	Ikaros zinc finger
11	676.5	60.1	470	1 W0970	Ikaros isoform i c
12	669	59.4	518	1 R2017	Murine Ikaros prot
13	669	59.4	518	1 W2674	Mouse Ikaros mix-1
14	669	59.4	518	1 W70966	Mouse Ikaros isofo
15	554	49.2	431	1 R20214	Murine Ikaros prot
16	551	48.9	431	1 R46963	Murine Ikaros zinc
17	551	48.9	431	1 W2671	Murine Ikaros mIk-
18	551	48.9	431	1 W70963	Mouse Ikaros isofo
19	432	38.4	432	1 R20216	Murine Ikaros prot
20	432	38.4	432	1 W2673	Mouse Ikaros mix-3
21	432	38.4	432	1 W70965	Mouse Ikaros isofo
22	418.5	37.2	334	1 R2020	Ikaros protein. Tr
23	418.5	37.2	334	1 W2677	Ikaros isoform. Ik
24	418.5	37.2	334	1 W70969	Human Ikaros prote
25	403.5	35.8	390	1 R20218	Murine Ikaros prot
26	403.5	35.8	390	1 W2675	Mouse Ikaros mix-4
27	403.5	35.8	390	1 W70967	Mouse Ikaros isofo
28	313	27.8	56	1 W2688	Ikaros protein SEQ
29	313	27.8	56	1 W2689	Ikaros protein SEQ
30	303	26.9	56	1 W2687	Ikaros protein SEQ
31	271.5	24.1	412	1 Y07088	Renal cancer assoc
32	269	23.9	711	1 W37504	Human OTK18. Novel
33	263	23.4	532	1 W52187	Human BRCA1-associ
34	258	22.9	706	1 R68743	BCL-6 zinc finger
35	258	22.9	181	1 R83490	Zif268-zif268 zinc
36	258	22.9	181	1 W84394	The zif268-zif268
37	256	22.7	561	1 W82404	Human SRE-ZBP anal
38	251.5	22.3	688	1 W06316	TRP-1 protein. DNA
39	249.5	22.2	803	1 W81756	Myc-binding zinc-f
40	244	21.7	421	1 W81631	Human stem cell zi
41	238.5	21.2	361	1 W81632	Human stem cell zi
42	230	20.4	88	1 W87557	A nucleic acid bin
43	230	20.4	88	1 W84299	An anti-HIV zinc i

44 230 20.4 88 1 W8390 zinc finger protel
45 227 20.2 500 1 W37948 Amino acid sequenc

ALIGNMENTS

RESULT 1
W15575
ID W15575 standard; Protein; 209 AA.
AC W15575
DT 07-JUL-1997 (first entry)
DE Human Aiolos polypeptide (partial sequence).
KW Aiolos; transcription activator; immune system; T lymphocyte;
KW B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW transgenic animal.
OS Homo sapiens.
PN W09714714-11.
PD 24-APR-1997.
PF 17-OCT-1996; U16774.
PR 18-OCT-1995; US-005529.
PR 14-MAY-1996; US-017646.
PA (GHEO) GEN HOSPITAL CORP.
PI Georgopoulos K, Morgan BA;
DR WPI; 97-245047/22.
DR N-PSDB; T60491.
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a mammalian immune system, for the treatment of T cell leukaemia(s) and lymphoma(s).
PS Disclosure; Fig 5b; 115pp; English.
CC A human Aiolos partial polypeptide (W15575) is encoded by an isolated cDNA clone (T60491). Aiolos is a homologue of Ikaros whose expression is restricted to lymphoid lineage. It forms homodimers that are potent transcriptional activators. It also forms less potent heterodimers with Ikaros. Recombinant Aiolos can be produced in transformed host cells. Aiolos misexpressing immune system components can be used to reconstitute the immune system of a mammal. The Aiolos protein, coding sequence or cells expressing CC Aiolos may be used to treat a disorder in an animal, esp. by gene CC therapy. Such disorders include T-cell leukaemias and lymphomas.
SQ Sequence 209 AA;

Query Match 100.0%; Score 1126; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.6e-105;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERDENVLKSEPMGNAEPEIPYSYSREYNEYENIKLERHVVSFDSSRPTSGKMNCDVCG 60
DB 1 ERDENVLKSEPMGNAEPEIPYSYSREYNEYENIKLERHVVSFDSSRPTSGKMNCDVCG 60

QY 61 SCISFNVLVHKRSHHTGERPQCNOCCGASTOKGNLRLHKLTGKPEKCHLCNVACOR 120
DB 61 SCISFNVLVHKRSHHTGERPQCNOCCGASTOKGNLRLHKLTGKPEKCHLCNVACOR 120

QY 121 RDALTGLHRTSHVEKPKCEFCGRYSKORSLEEHKERCRTFLQSTDPGTASAEARHK 180
DB 121 RDALTGLHRTSHVEKPKCEFCGRYSKORSLEEHKERCRTFLQSTDPGTASAEARHK 180

QY 181 AEMGSEALVLDRLASNVAKRKSMPQKF 209
DB 181 AEMGSEALVLDRLASNVAKRKSMPQKF 209

RESULT 2
W15574
ID W15574 standard; Protein; 507 AA.
AC W15574;
DT 07-JUL-1997 (first entry)
DE Mouse Aiolos polypeptide.
KW Aiolos; transcription activator; immune system; T lymphocyte;
KW B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW transgenic animal.

THIS PAGE BLANK (USPTO)

LOCUS AA080413 339 bp mRNA EST 19-FEB-1998
 DEFINITION ca43h06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1307771 3' similar to TR:000598 000598 HIK1.; mRNA sequence.
 ACCESSION AA080413
 NID 92877819
 VERSION AA080413.1 GI:2877819
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 339)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2153413.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 457 Std Error: 0.00

Seq primer: -40m13 fwd. Et from Amersham

High quality sequence stop: 295.

Location/Qualifiers

FEATURES

source

1. 339

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCI_CGAP_GCB1"

/tissue_type="germinal center B cell"

/lab_host="DH10B"

/note="Vector: pT7T30-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was

primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCATTTTTTTTTTTTTTTT-

3']. Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

69 a 102 c 75 g 93 t

BASE COUNT

ORIGIN

Query Match 14.8%; Score 93; DB 38; Length 339;

Best Local Similarity 72.7%; Pred. No. 5.1e-18;

Matches 120; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 140 ggccaaccagtggaagatgaactgcgatgtgtgtggattatcctgcagcttcaatg 199

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 169 GACCTTCTAACGAAACAAAGTGTGATATCTGTGGGATCATTTGCATCGGCCCATG 110

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 200 tcttaatggttcataagcgaaccatactgtgtgaacgccccattccagtgtaactagtgtg 259

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 109 TGCTCATGTTTCACAAAAGACACACTGGAGAGCGGCCCTTCCAGTCAATCAGTGG 50

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 260 gggcatctttactcagaaggttaacctctctccgcacattaaac 304

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 49 GGGCCTCATTACCCAGAGAGGCAACCTGCTCCGGCACAATCAAGC 5

RESULT 7

LOCUS R19256

DEFINITION Y924h05.r1 Soares infant brain INIB Homo sapiens cDNA clone

IMAGE:33351 5', mRNA sequence.

ACCESSION R19256

NID 9772866

VERSION R19256.1 GI:772866

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 467)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1113

High quality sequence stops: 311 Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1113 Std Error: 0.00

Seq primer: M13Rp1

High quality sequence stop: 311.

FEATURES

source

1. 467

/organism="Homo sapiens"

/db_xref="GDB:405698"

/db_xref="taxon:9606"

/clone_lib="IMAGE:33351"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Laimid BA; Site_1: Not

I; Site_2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5'

AACTGGAAGAAATTCGCGCGCGAGGAAATTTTTTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the Laimid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo."

138 a 84 c 124 g 117 t 4 others

BASE COUNT

ORIGIN

Query Match 14.3%; Score 90; DB 21; Length 467;

Best Local Similarity 55.9%; Pred. No. 4.6e-17;

Matches 171; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 163 tgcgatgtgtgattatcctgcagcttcatttccttaagtgttcataagcgaagc 222

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 2 TGTAATGAGTGTGGAAGAGCGCTTCAGCGTTAATGGGAACTAATCGGCATCAGAGATT 61

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 223 catactggtgaacgccccattccagtgtaatacagtggtgggcattctttactcagaaggt 282

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 62 CACACTGGGGAGAAACCTTTTGAATGTAATGAGTGTGGGAGATGCTTTTACTTCTAAAGA 121

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:673927
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 423.

FEATURES

Location/Qualifiers
 1. .609

source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:128127"
 /clone_lib="Soares 2NDMT"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGGAGCGCGGTTTGTGTGATTATCTGATGCTTCTT
 3']]; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."

BASE COUNT 124 a 162 c 152 g 171 t

ORIGIN

Query Match 28.6%; Score 179.6; DB 40; Length 609;
 Best Local Similarity 87.8%; Pred. No. 3e-44;
 Matches 230; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

QY 138 caggccaacagtggaagatgaactgcgatgtgtgtgattatctctgcatcagcttc-a 196

Db 595 CAGACCAACCGGTGGAAGTAACATGCGACGTGCGGTTATCCCTGCATTAGCTTCAA 536

QY 197 atgtcttaatggttcataagcgaagccatactggtgaacgccattccagtgtaacacg 255

Db 535 ACCTCTTGATGTTCTAFAAGCGAAGCCATACCCGCGAGCCCGTTCACAGTGTATCAG 476

QY 256 tgtgggcatctttactcagaaggaagtaacctctccgcacattataaactgcacacagg 315

Db 475 TG-CGGGCATCTTTTACTCAGAAAGTAACCTCTCCGTCAATATTAACTGCACACGGGG 417

QY 316 gaaaaacatttaagtgtcacctctgcaactatgcatgcaagaagagatgcgctcacg 375

Db 416 GAAAAACCTTTAAAGTGTCACTCTGCAACTAGCATGCGTCAAGAGAGATGCGCTCAGG 357

QY 376 gggcatcttaggacacattctg 397

Db 356 GGACACCTTAGGACACATCTG 335

RESULT

AI325349/c 5
 LOCUS AI325349 557 bp mRNA EST 23-DEC-1998
 DEFINITION mi06e03.y1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:459676 5' similar to gb:L03547 Mouse Ikaro DNA binding
 protein (MOUSE);, mRNA sequence.

ACCESSION AI325349

NID 94059778

VERSION AI325349.1 GI:4059778

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 557)

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE

The WashU-HMT Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

On Jan 17, 1998 this sequence version replaced gi:1900564.

Contact: Marra M/Mouse EST Project

WashU-HMT Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:276564

This read is a RESEQUENCE of a previously sequenced mouse clone

This read has been verified (found to hit its original self in the

correct orientation)

Seq primer: -40RP from Gibco

High quality sequence stop: 417.

Location/Qualifiers

1. .557

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/map="21"

/clone_image="IMAGE:459676"

/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"

/sex="unknown"

/tissue_type="placenta"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)

with a modified polylinker; Site_1: Not I; Site_2: Eco RI;

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 116 a 157 c 127 g 155 t 2 others

ORIGIN

Query Match 19.9%; Score 125; DB 45; Length 557;

Best Local Similarity 75.6%; Pred. NO. 1.1e-27;

Matches 155; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcatcagcttcaatgcttaa 205

Db 206 CTAACGGAAAACTAAAGTGTGATATCTGTGGATCGTTTGCATCGGCGCCCAATGTGCTCA 147

QY 206 tggttcataagcgaagccatactggtgaacgccattccagtgtaacagtggtgggcat 265

Db 146 TGGTTTCAAAAAGAGTGTATCTGTGTGAACGCCCTTTCAGTGCACACCATGTTGGGGCT 87

QY 266 cttttactcagaaagtgtaacctctccgccacattaaactgcacacaggggaaaaacctt 325

Db 86 CCTTTACCAGAAAGCAACCTCTCTGCGGCACATCAACGTGCACCTCGGTGAGAGACCT 27

QY 326 ttaagtgtcacctctgcacattatgc 350

Db 26 TCAAAATGCCATCTTTGCAACTATGC 2

RESULT

6

AA808413/c

```
Query Match      38.2% Score 239.6; DB 48; Length 546;
Best Local Similarity 85.3%; Pred. No. 1.7e-62;
Matches 279; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 1 gaaagagatgagaatgttttaagtcagaaacccatggagaaatgcagagagccctgaaatc 60
DB 223 GACAGAGATGAGAACATATGAAACCGGAGCCCATGGAGATGAGAGAGAGTGAATG 282
QY 61 ccttcacagctattcaagagaataataagataaataaataagttggagagacatgtt 120
DB 283 CCTTACAGCTATGACAGAGATAGACGACTATGAAGCATTAGCTGGAGAGACA---C 339
QY 121 gtctcaatcgatagtagcaggcccaaccagtggaagatgaactcgatgtgtggtatta 180
DB 340 GTGCCCTATGACACAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAG 399
QY 181 tctgcagctattcaagagaataataagataaataaataagttggagagacatgtt 240
DB 460 TTCCAGTGTATGATGAGTGGGGGCGATCTTTTACTCAGAAAGTAACTCCCTCGTCATATT 519
QY 301 aaactgcacacaggggaaaaaccccttt 327
DB 520 AAACGTGCACACGGGAAAAAACCTTTT 546

RESULT 3
AI550402      501 bp mRNA EST 23-MAR-1999
LOCUS      vx13e05.x1 Soares 2NMT Mus musculus cDNA clone IMAGE:1264352 3'
DEFINITION      similar to TR:008900 008900 AIOLOS ; mRNA sequence.
ACCESSION      AI550402
NID      94482765
VERSION      AI550402.1 GI:4482765
KEYWORDS      EST.
SOURCE      Mus musculus
ORGANISM      house mouse.
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Marra,M., Hallier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced gi:3246999.

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:666904
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
High quality sequence stop: 381.
Location/Qualifiers
1. 501
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

FEATURES
source
```

Qy	456	tgaggagcacaagagcgctgcgttacattcttcagagcactgaccacagggacactgc	515
Db	338	GGAGGAGCACAAGGAAGCGCTGCGGAGCTTTTCTTCAGAACCCCTCACCTGGGGACGCTGC	279
Qy	516	aagtgcgagcgcaagcacacatcaaaacagagatggaagtgaagagctctcgactcga	575
Db	278	AAGTGTGGAGGCAAGACACATCAAAAGCCGAGATGGGAAGTGAGAGAGCTCTCGTCTTGA	219
Qy	576	cagattagcaagcaattgttgcaaaacgaaaaaagctcaatgcctcagaaattca	628
Db	218	CAGATTAGCAAGCAATGTGCTTAAGCGAAAAGCTCGATGCTCAGAAATCA	166

RESULT

2

AI549961

LOCUS

DEFINITION

ACCESSION

NID

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI549961

546 bp mRNA

23-MAR-1999

vx02e08.x1 Soares 2NDMT Mus musculus cDNA clone IMAGE:1263302

similar to TR:008900 O08900 AIOLOS ; mRNA sequence.

AI549961

94482324

AI549961.1 GI:4482324

EST

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 546)

Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Rensing,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

On Jun 5, 1998 this sequence version replaced gi:3189635.

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:665854

This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyt not found

High quality sequence stop: 460.

FEATURES

SOURCE

1..546

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:1263302"

/clone_lib="Soares 2NDMT"

/sex="male"

/tissue_type="Thymus"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker: Site.1: Site.2: Eco RI: 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTAACATCTGAAGTGGAGCGCGGCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

173 a 129 c 149 g 94 t 1 others

BASE COUNT

ORIGIN

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 08:16:48 ; Search time 437.25 Seconds
(without alignments)
2833.050 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagagagaaatgtttt.....ctcaatgcctcagaaattca 628

Scoring table: IDENTITY_NUC
Searched: 2546578 seqs, 986266752 residues

Database : EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

54: em_est22:*

55: em_est23:*

56: em_est24:*

57: em_est25:*

58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	303.4	48.3	577	39	AA863970	AA863970 vx87d08.r
C 2	239.6	38.2	546	48	A1549961	A1549961 vx02e08.x
C 3	195.5	29.5	501	48	A1550402	A1550402 vx13e05.x
C 4	179.6	28.6	609	40	AA920812	AA920812 vx86f12.r
C 5	125.9	19.9	557	45	A1325349	A1325349 m106e03.y
C 6	93.3	14.8	339	38	AA808413	AA808413 oa43h06.s
C 7	90.0	14.3	467	21	R19256	R19256 y924h05.r1
C 8	89.6	14.3	355	24	H83408	H83408 y90h12.r1
C 9	89.4	14.2	509	27	AA027561	AA027561 m106e03.r
C 10	85.6	13.6	291	32	AA334780	AA334780 EST39041
C 11	85.2	13.6	470	21	R17304	R17304 y912d09.r1
C 12	85.2	13.6	431	22	R22824	R22824 y947a01.r1
C 13	84.6	13.5	389	28	AA119538	AA119538 mp68a03.r
C 14	84.4	13.4	587	48	A1559335	A1559335 tg43h01.x
C 15	83.8	13.3	382	33	AA400326	AA400326 zu69b10.r
C 16	83.6	13.3	525	29	AA130717	AA130717 zol13h06.s
C 17	83.4	13.3	644	22	H05551	H05551 y175b08.r1
C 18	83.3	13.2	805	49	A1648952	A1648952 uk33c08.x
C 19	82.8	13.2	332	29	AA171117	AA171117 ms55b10.r
C 20	82.8	13.2	512	46	A1406535	A1406535 EST23481
C 21	82.8	13.2	406	54	HSW009139	HSW009139 Homo sapi
C 22	82.6	13.2	837	53	HSW007817	HSW007817 Homo sapi
C 23	82.1	13.1	494	39	AA848430	AA848430 EST191190
C 24	81.8	13.0	704	53	HSW007214	HSW007214 Homo sapi
C 25	80.8	12.9	645	36	AA625515	AA625515 af72e07.r
C 26	80.8	12.9	544	44	A1243713	A1243713 qb88f01.x
C 27	80.2	12.8	758	36	AA628972	AA628972 af28h12.s
C 28	80.2	12.8	791	38	AA772346	AA772346 a143c08.s
C 29	80.2	12.8	768	38	AA789218	AA789218 aj27c09.s
C 30	80.2	12.8	723	40	AA972741	AA972741 op90h05.s
C 31	79.8	12.7	424	26	W19169	W19169 zb25b08.r1
C 32	79.8	12.7	369	39	AA814418	AA814418 nz12d08.s
C 33	79.6	12.7	601	34	AA455712	AA455712 aa22e04.r
C 34	79.2	12.6	774	36	AA628881	AA628881 af30a12.s
C 35	78.8	12.5	662	26	W28371	W28371 46a4 Human
C 36	78.8	12.5	491	35	AA548358	AA548358 nk41a05.s
C 37	78.8	12.5	545	36	AA624117	AA624117 vg75g07.s
C 38	78.6	12.5	447	23	R98368	R98368 HH210SF Hum
C 39	78.6	12.5	447	23	R98368	R98368 HH211SF Hum
C 40	78.6	12.5	761	40	AA917951	AA917951 ol168g11.s
C 41	78.4	12.5	685	26	W27601	W27601 35a3 Human
C 42	78.4	12.5	474	28	AA113952	AA113952 zm81c11.r
C 43	78.4	12.5	465	42	A1128612	A1128612 ga62a10.s
C 44	78.4	12.5	492	44	A1311718	A1311718 g092f08.x
C 45	78.2	12.5	378	21	T63663	T63663 yc23a03.r1

ALIGNMENTS

RESULT 1
AA863970/c
LOCUS vx87d08.r1 Soares 2Nbmt Mus musculus cDNA clone IMAGE:1282191 5'
DEFINITION similar to TR:008900 008900 AIOLOS ;, mRNA sequence.
ACCESSION AA863970
NID 92956449
VERSION AA863970.1 GI:2956449


```

RESULT 15
US-08-933-750C-53
: Sequence 53, Application US/08933750C
: Patent No. 5932442
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Purvi
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals
: STREET: 3174 Porter Drive
: CITY: Palo Alto

```

```

Db 1759 GGGAGGACCTTAAATGATCGCTCAGGCCATCTCCCCAGCACCTTGAGGACTTCACACTGGCGCT 1811
Qy 403 aaacctacaaaatgtgagtttctgtgaagggagttacaaagcagag 446
    || ||||| || || ||||| || |||||
Db 1819 AAGCCCTACAAGTGTCAAGGACTGTGAAAAGCCTTCCGCCCAGAG 1862

```

Search completed: November 6, 1999, 06:57:40
Job time: 141 sec

Search completed: November 6, 1999, 06:57:40
Job time: 141 sec


```

1 ZIP: 02109
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 COMPUTER: IBM PC compatible
6 OPERATING SYSTEM: PC-DOS/MS-DOS
7 SOFTWARE: ASCII (text)
8
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: PCT/US95/09345
11 FILING DATE:
12
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 08/283,300
15 FILING DATE: 29-JULY-94
16
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 08/238,212
19 FILING DATE: 03-MAY-94
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 08/121,438
23 FILING DATE: 14-SEP-1993
24
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 07/946,233
27 FILING DATE: 14-SEP-1992
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Myers, Paul L.
30
31 REGISTRATION NUMBER: 35,695
32 REFERENCE/DOCKET NUMBER: MGP-027PC
33
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (617)227-7400
36 TELEFAX: (617)227-5941
37
38 INFORMATION FOR SEQ ID NO: 3:
39
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 1296 base pairs
42 TYPE: nucleic acid
43 STRANDEDNESS: single
44 TOPOLOGY: linear
45 MOLECULE TYPE: CDNA
46 FEATURE:
47 NAME/KEY: CDS
48 LOCATION: 1..1296
49
50 PCT-US95-09345-3

```

Query Match	23.7%	Score 148.6;	DB 5;	Length 1296;
Best Local Similarity	73.4%	Pred. No. 3e-39;		
Matches 190: Conservative	0;	Mismatches 69;	Indels 0;	Gaps 0;

Qy	146	ccagtggaagatgaactgcgagtgtgtgtgattatctcgatcgatcagcttcgaatgtcttaa	205
Db	338	CTAACGGAAACTAAGTGTGATATCTGTGGATCTGTTTCATCGGGCCCAATGTGCTCA	397
Qy	206	tggttcataagcgaagccatactggtggaagcccatccagtgtaatcagtgtagggcat	265
Db	398	TGGTTTCAAAAAGAGTCATCTGTGTGAACGGCTTTCAGTGCACCACTCTGGGGCT	457
Qy	266	cttttactcagaaggttaacctctctcgccacataaactgcacaggggaaaaacctt	325
Db	458	CCTTTACCCGAAAGGCAACTCTCTCGGSCACATCAAGCTGCCTCGGTGAGAAGCCCT	517
Qy	326	ttaagtgtaacctctcaactatgcatgcgaagaagagatcgctcaccgggcatctta	385
Db	518	TCAATGCCATCTTTGCAACTATGCTCGCGCGGAGGACGCCCTTACCGGCCACCTGTA	577
Qy	386	ggacacattctgtggagaa	404
Db	578	GGACGCATCCGGAGACAA	596

```

RESULT 11
US-08-465590-6
: Sequence 6, Application US/08465590
: Patent No. 5824770
: GENERAL INFORMATION:
: APPLICANT: Georgopoulos, Katia A.
: TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE

```

```

NUMBER OF SEQUENCES: 164
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Ascii (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: \US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MPG-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1170 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1170
US-08-465,590-6

```

Query Match	17.1%	Score 107.6;	DB 3;	Length 1170;
Best Local Similarity	77.1%	Pred. No. 7.6e-26;		
Matches 131: Conservative	0;	Mismatches 39;	Indels 0;	Gaps 0;

229	QY	ggtgaacgccattccagtgtaatacagtggtgggcatcttttactcagaaaaggtaacctc	288
160	Db	GGTGAACGGCCITTCAGTGCACCAAGTCTGGGCTCCTTTACCCAGAAAGCAACCTC	219
289	QY	ctccggcaccattaaactgcacacaggggaaaaaccttttaagtgtccactctgcacatat	348
220	Db	CTGGGCACATCAAGTGCACCTCGGTGAGAGGCCCTTCAATGCACCTTTTGGCACTAT	279
349	QY	gcattgccaagaagagatgcgtccacggggcatctctaggacacattcttgt	398
280	Db	GGCTGGCCGGGAGGAGCCCTCTCACGGCCCACTTGAGGACGCACTCCCT	329

RESULT 12
PCT-US95-09345-5
; Sequence 5, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON

NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,695
 REFERENCE/DOCKET NUMBER: MGP-027PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1788 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 223..1515
 PCT-US95-09345-1

Query Match 27.1%; Score 170; DB 5; Length 1788;
 Best Local Similarity 67.1%; Pred. No. 3.5e-46;
 Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;
 QY 229 ggtgaagccattccagtgtaataatcagtggtgggcatcttttactcagaaggttaacctc 288
 Db 382 GGTGAAGCGGCTTTCCAGTGCACCACTGTGGGGCCTTCCCTTACCCAGAAAGGCAACCTC 441
 QY 289 ctccgcacattaaactgcacacaggggaaacaccttttaagtgtcacctctgcaactat 348
 Db 442 CTGCGGCACATCAAGCTGCCTCGGTGAGAGCCCTTCAATGCCATCTTTGCAACTAT 501
 QY 349 gcatgccaaagaagatgcgtccacggggcatctttaggacacattctgtgggaaaccc 408
 Db 502 GCCTGCCGCGGGAGGAGCGCCCTCACCGGCACCTGAGGAGCGCACTCGCTTGGTAAGCCT 561
 QY 409 tacaattgtagtttttgaagagagtataacagcagaagttcccttgaggagcacaag 468
 Db 562 CACAAATGTGATATTGTGGCGGAGGTATAAACAGCGGAAGCTTTTAGAGGAGCATAAA 621
 QY 469 gagcgctgcgtacattcttcagagcactgacc-----caggggacactgcaagtgcg 522
 Db 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCGGGCGGTGTCGCCAGTCATTAA 681
 QY 523 gaggcagacacataaag-----cagagatgggaagtgaagagct 564
 Db 682 GAAGAAACTTACCACACAGAGATGGCAGAGACCTGTGCAAGATAGGAGCAGAGAGGTCC 741
 QY 565 ctcttactgaagatagcaagcaatgtggcaaaacgaaacacacacacacacacacacac 624
 Db 742 CTTCCTGGACAGGCTGGCAAGCAATGTGCGCAACGTAAGAGCTCTATGCCTCAGAAA 801
 QY 625 tt 626
 Db 802 TT 803

RESULT 9
 US-08-465-590-4
 : Sequence 4, Application US/08465590
 : Patent No. 5824770
 : GENERAL INFORMATION:
 : APPLICANT: Georgopoulos, Katia A.
 : TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
 : NUMBER OF SEQUENCES: 164
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: LAHIVE & COCKFIELD
 : STREET: 60 STATE STREET, Suite 510
 : CITY: BOSTON
 : STATE: MASSACHUSETTS
 : COUNTRY: USA
 : ZIP: 02109
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII (text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465,590
 FILING DATE: 05-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/238,212
 FILING DATE: 02-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/121,438
 FILING DATE: 14-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/946,233
 FILING DATE: 14-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,695
 REFERENCE/DOCKET NUMBER: MGP-006C2DV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1296 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1296
 US-08-465-590-4

Query Match 23.7%; Score 148.6; DB 3; Length 1296;
 Best Local Similarity 73.4%; Pred. No. 3e-39;
 Matches 190; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
 QY 146 ccagtggaagatgaactgcgatgtgtggattatctctgcacatcagcttcaatgtcttaa 205
 Db 338 CTAAACGAAACTAAAGTGTGATATCTGGGATCTGTTGCAATCGGGCCCAATGTGCTCA 397
 QY 206 tggttcataagcgaagcactactgtgtgaacgccattccagtgtaatcagtggtgggcat 265
 Db 398 TGGTTCACAAAGAAAGTCATACCTGGTGAACGGCCTTTCCAGTGCACCACTCTGGGGCCT 457
 QY 266 ctcttactcagaaggttaacctctctcgccacattaaactgaacagaggggaaaaacctt 325
 Db 458 CCTTACCAGAAAGGCAACCTCTCGCGCACATCAAGCTGCACTCGGTGTGAGAAGCCCT 517
 QY 326 ttaagtgtcacctctgcaactatgcatgccaagaagagagatgcgtcacggtgcatctta 385
 Db 518 TCAATGCCATCTTGCACACTATGCTCGCGCACATCAAGCTGCACTCGGTGTGAGAAGCCCT 577
 QY 386 ggacacattctgtggagaa 404
 Db 578 GGACGCACTCCGGAGACAA 596

RESULT 10
 PCT-US95-09345-3
 : Sequence 3, Application PC/TUS9509345
 : GENERAL INFORMATION:
 : APPLICANT: Georgopoulos, Katia A.
 : TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
 : NUMBER OF SEQUENCES: 26
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: LAHIVE & COCKFIELD
 : STREET: 60 STATE STREET, Suite 510
 : CITY: BOSTON
 : STATE: MASSACHUSETTS
 : COUNTRY: USA

2:

```

RESULT 7
PCT-US93-08743-2
; Sequence 2, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1788 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 223..1515
PCT-US93-08743-2

```

Qy	229	ggtgaagccccaattcaagtgttaactgaatgaatgtgtggggacatcttttactcagaaagtgaaacctc	288
Db	382	GGTGAACGGGCTTTCCAGTGCACACCGTCTGGGGCCTCTTTTACCAGAAAGGCAACCTC	441
Qy	289	ctcgcacacattaaactgcacacaggggaaaaacctttaagtgtcacctctcgaactat	348
Db	442	CTGGGGCACATCAAGCTGCACACTCGGTGAGAACCCCTTCAATGCCACTCTTTGCCAACTAT	501
Qy	349	gcatacgcaagaagagatgcgtccacgggggcatcttaggacacattctgtgggagaaaccc	408
Db	502	GCCTGCCGGCGGAGGACGCCCTCACCGGCCACCTGAGGACGCACCTCCGTGTGTAAGCCT	561
Qy	409	tacaaatgtgagttttgtggaagaggttacaaacagagagaagtcccttggaggagacaaag	468
Db	562	CACAAATGTGGATATTGTGCGCCGGAGCTATAACACGCGAAGCTCTTTAGAGGAGCATAA	621
Qy	469	gagcgtgcggtacattcttcagagcactgacc-----caggggacacatgcgaagtgcg	522
Db	622	GAGCGATGCCACAACTACTTTGGAAGCATGGGCTTCCGGGGCTGTGCCAGTCATTAG	681
Qy	523	gaggcaagacacatcaaaag-----cagagatgggaagtgaagaagct	584
Db	682	GAAGAAACTTAACCAACACGAGATGGCAGAGAGACTGTGTCAAGATAGGAGCAGAGAGTCC	741
Qy	565	ctcgtactgacagattagcaagcaatgtggcaaaaacgaaaaagctcaatgcctcagaaa	624
Db	742	CTTGTCTCTGACAGGCTGGGCACCAATGTCCGCCAAACGTTAAGAGCTCTATGCCCTCAGAAA	801
Qy	625	tt 626	
Db	802	tt 803	

```

RESULT      8
PCI-US95-09345-1
; Sequence 1, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:

```

Db 1040 GTAAGAGCTCTATGCTCAGAAATT 1064

RESULT 5

PCT-US95-09345-4

Sequence 4, Application PC/TUS9509345

GENERAL INFORMATION:

APPLICANT: Georgopoulos, Katia A.

TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 STATE STREET, Suite 510

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/09345

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/283,300

FILING DATE: 29-JULY-94

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/238,212

FILING DATE: 02-MAY-94

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/121,438

FILING DATE: 14-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/946,233

FILING DATE: 14-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Paul L.

REGISTRATION NUMBER: 35,695

REFERENCE/DOCKET NUMBER: MGP-027PC

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2049 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 223..1776

PCT-US95-09345-4

Query Match 33.4%; Score 209.8; DB 5; Length 2049;

Best Local Similarity 67.1%; Pred. No. 3.4e-59;

Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;

Qy 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcacatcagcttcaatgtcttaa 205

Db 560 CTACGGAAACCTAAAGTATCTGTGGGATCGTTTGCATCGGGCCCAATGTGCTCA 619

Qy 206 tgggtcaagcgaagccactactggtgaacgccattccagttcaatcagtggtggcatt 265

Db 620 TGGTTCACAAAGAGTCATCTGTTGGAACGGCTTTCCAGTGCACCACTCTGGGGCCT 679

Qy 266 cttttactcagaaggttaacctctccgcacacatttaactgcacacaggggaaaaacctt 325

Db 680 CCTTTACCCGAAAGCAACCTCTCTCGGGCACATCAAGCTGACTCGGGTGAGAGCCCT 739

Qy 326 ttaagtgcaactctgcacactatgcacacaaagagatgcgctcaccgggggcatcttta 385

Db 740 TCAAAATGCCATCTTTTGAACATATGCTCCGCGGAGGAGCGCCCTCACCGGCCACCTGA 799

Qy 386 gacacattctgtgagaaacccctacaatgtgagttttgtggaaggagttlacaagcaga 445

Db 800 GGACCGCACTCCGTGGTAAGCTTCAAAATGTGGNATNTGTGGCCGGAGCTATAAACAGC 859

Qy 446 gaagttcccttgagagcacaagagcgctgccgtacatttcttcagagcactgacc--- 503

Db 860 GAAGCTCTTTAGAGGAGCATAAAGAGCGATGCCAACACTACTTGGAAAGCATGGGCCCTTC 919

Qy 503 ---caggggacactgcaagtcgagggaggaagacacacatcaag----- 542

Db 920 CGGCGCTGTGCCAGTCATTAAAGGAAGAAACTAACCAACAGATGCGCAGAAGACCTGT 979

Qy 542 cagagatgggaagtgaagagctctcgtactggacagattagcaagcaatgtggcaaac 601

Db 980 GCAAGATAGGACGACAGAGGTCCTTGTCTCTGGACAGCTGGCAAGCAATGTGCGCAAC 1039

Qy 602 gaaaagctcaatgcctcagaatt 626

Db 1040 GTAAGAGCTCTATGCTCAGAAATT 1064

RESULT 6

US-08-465-590-2

Sequence 2, Application US/08465590

Patent No. 5824770

GENERAL INFORMATION:

APPLICANT: Georgopoulos, Katia A.

TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE

NUMBER OF SEQUENCES: 164

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 STATE STREET, Suite 510

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/465,590

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/238,212

FILING DATE: 02-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/121,438

FILING DATE: 14-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/946,233

FILING DATE: 14-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Paul L.

REGISTRATION NUMBER: 35,695

REFERENCE/DOCKET NUMBER: MGP-006C2DV

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1788 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

QY 206 tggttcataagcgaagccatactagtggaacgcccattccagtgtaatacagtggtgggcat 265
 Db 233 TGGTTACAAAAGAACCCACTGGAGAACGGCCCTTCAGTGCATCAGTGGGGCCT 292
 QY 266 cttttactcagaaggttaacctctccgacattaaactgcacacaggggaaaaacctt 325
 Db 293 CATTCACCCAGAGGCAACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 352
 QY 326 ttaagtgtcacctctgcaactatgcacaaagaagagatgcgtcacggggcatctta 385
 Db 353 TCAATATGCCACCTCTGCAACTACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 412
 QY 386 ggaacacattctgtgagaacacctcaaatgtgagtttggagagattacaagcaga 445
 Db 413 GGACGACCTCCGTTGTAACCTCAACAATGTGGATATTGTGGCCGAAGCTATAAACAGC 472
 QY 446 gaagttcccttgaggagcacaagggagcgctgcgtacatttcttcagagcactgacc--c 503
 Db 473 GAACGCTTTAGAGGAACATTAAGAGCGCTGCCACAACCTACTTGGGAAGCATGGGCCCTTC 532
 QY 504 aggggacactgcaagtgcggagggaagacacacatcaaaagca----- 544
 Db 533 CGGGCACACTGACCCAGTCATTAAGAGAACTAAGCACAGTGAATGGCAGAGACC 592
 QY 544 -----gagatgggaagtgaagagctctctgactggacagattagcaacaatgtggcaa 598
 Db 593 TGTGCAAGATAGGATCAGAGAGATCTCTGCTGTCGACAGACTAGCAAGTAATGTGCGCA 652
 QY 599 aacgaaaaagctcaatgcctcagaaatt 626
 Db 653 AAGTAAGAGCTATGCTTCAGAAATT 680

RESULT 2

PCT-US93-08743-3
 ; Sequence 3, Application PC/TUS9308743
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
 ; NUMBER OF SEQUENCES: 152
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/08743
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 946,233
 ; FILING DATE: 14-SEP-1992
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1611 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1611
 ; PCT-US93-08743-3

Query Match 34.5%; Score 216.4; DB 5; Length 1611;
 Best Local Similarity 67.9%; Pred. No. 2.1e-61;
 Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

QY 146 ccagtggaagatgaactgcgagtgtgtgattatctcgtcatcagctcaatgtcttaa 205
 Db 395 CTAACGGAAACTAAAGTGTGATCTGTGGGATCATTTGATCGGGCCCAATGTGCTCA 454

QY 206 tggttcataagcgaagccatactagtggaacgcccattccagtgtaatacagtggtgggcat 265
 Db 455 TGGTTACAAAAGAACCCACTGGAGAACGGCCCTTCAGTGCATCAGTGGGGCCT 514
 QY 266 cttttactcagaaggttaacctctccgacattaaactgcacacaggggaaaaacctt 325
 Db 515 CATTCACCCAGAGGCAACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 574
 QY 326 ttaagtgtcacctctgcaactatgcacaaagaagagatgcgtcacggggcatctta 385
 Db 575 TCAATATGCCACCTCTGCAACTACCTGCTCCGGCACAATCAAGTGTGATTCGGGGGAGAAGCCCT 634
 QY 386 ggaacacattctgtgagaacacctcaaatgtgagtttggagagattacaagcaga 445
 Db 635 GGACGACCTCCGTTGTAACCTCAACAATGTGGATATTGTGGCCGAAGCTATAAACAGC 694
 QY 446 gaagttcccttgaggagcacaagggagcgctgcgtacatttcttcagagcactgacc--c 503
 Db 695 GAACGCTTTAGAGGAACATTAAGAGCGCTGCCACAACCTACTTGGGAAGCATGGGCCCTTC 754
 QY 504 aggggacactgcaagtgcggagggaagacacacatcaaaagca----- 544
 Db 755 CGGGCACACTGACCCAGTCATTAAGAGAACTAAGCACAGTGAATGGCAGAGACC 814
 QY 544 -----gagatgggaagtgaagagctctctgactggacagattagcaacaatgtggcaa 598
 Db 815 TGTGCAAGATAGGATCAGAGAGATCTCTGCTGTCGACAGACTAGCAAGTAATGTGCGCA 874
 QY 599 aacgaaaaagctcaatgcctcagaaatt 626
 Db 875 AAGTAAGAGCTATGCTTCAGAAATT 902

RESULT 3

PCT-US95-09345-2
 ; Sequence 2, Application PC/TUS9509345
 ; GENERAL INFORMATION:
 ; APPLICANT: Georgopoulos, Katia A.
 ; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 STATE STREET, Suite 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII (text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/09345
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/283,300
 ; FILING DATE: 29-JULY-94
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/238,212
 ; FILING DATE: 02-MAY-94
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/121,438
 ; FILING DATE: 14-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/946,233
 ; FILING DATE: 14-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Paul L.
 ; REGISTRATION NUMBER: 35,695
 ; REFERENCE/DOCKET NUMBER: MGP-027PC
 ; TELECOMMUNICATION INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 06:57:37 ; Search time 44.29 Seconds
(without alignments)
1475.260 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagatgagaattttt.....ctcaatgctcagaattca 628

Scoring table: IDENTITY_NUC

Searched: 192659 seqs, 52021692 residues

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS9_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216.4	34.5	1386	3	US-08-465-590-3
2	216.4	34.5	1611	5	PCT-US93-08743-3
3	216.4	34.5	1386	5	PCT-US95-09345-2
4	209.8	33.4	2049	3	US-08-465-590-5
5	209.8	33.4	2049	5	PCT-US95-09345-4
6	170	27.1	1788	3	US-08-465-590-2
7	170	27.1	1788	5	PCT-US93-08743-2
8	170	27.1	1788	5	PCT-US95-09345-1
9	148.6	23.7	1296	3	US-08-465-590-4
10	148.6	23.7	1296	5	PCT-US95-09345-3
11	107.6	17.1	1170	3	US-08-465-590-6
12	107.6	17.1	1170	5	PCT-US95-09345-5
13	106.6	17.0	1004	3	US-08-465-590-8
14	106.6	17.0	1004	5	PCT-US95-09345-7
15	72.8	11.6	2031	4	US-08-933-750C-53
16	69.6	11.1	1309	4	US-08-933-750C-63
17	64.8	10.3	1892	4	US-08-933-750C-66
18	62	9.9	2133	3	US-08-820-170A-11
19	62	9.9	3754	3	US-08-820-170A-12
20	58.4	9.3	4252	5	PCT-US95-08429-4
21	56.6	9.0	3720	1	US-08-074-967-1
22	56.6	9.0	3720	4	US-08-553-541B-1
23	56.6	9.0	3720	5	PCT-US94-05669-1
24	55.2	8.8	3810	3	PCT-US95-08429-8
25	49.2	7.8	267	3	US-08-040-548-14
26	49.2	7.8	3086	3	US-08-040-548-15
27	49.2	7.8	267	3	US-08-466-344-14
28	49.2	7.8	3086	3	US-08-466-344-15
29	49.2	7.8	3068	3	US-08-224-482-1
30	48	7.6	2043	4	US-08-388-590A-39
31	48	7.6	5648	5	PCT-US96-03940-1
32	48	7.6	756	5	PCT-US96-03940-2
33	48	7.6	1407	5	PCT-US96-03940-3
34	48	7.6	1090	5	PCT-US96-03940-4
35	48	7.6	928	5	PCT-US96-03940-5
36	48	7.6	1791	5	PCT-US96-03940-6
37	48	7.6	3291	5	PCT-US96-03940-10

38 48 7.6 1461 5 PCT-US96-03940-27 Sequence 27, Appl
39 48 7.6 1284 5 PCT-US96-03940-28 Sequence 28, Appl
40 47.6 7.6 3132 3 US-08-224-482-3 Sequence 3, Appl
41 46.8 7.5 1161 4 US-08-616-857-1 Sequence 1, Appl
42 46.6 7.4 1420 4 US-08-909-965C-3 Sequence 3, Appl
43 46.4 7.4 2811 3 US-08-040-548-31 Sequence 31, Appl
44 46.4 7.4 2811 3 US-08-466-344-31 Sequence 31, Appl
45 46.4 7.4 1440 3 US-08-224-482-5 Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-465-590-3
: Sequence 3, Application US/08465590
: Patent No. 5824770
: GENERAL INFORMATION:
: APPLICANT: Georgopoulos, Katia A.
: TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
: NUMBER OF SEQUENCES: 164
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 STATE STREET, Suite 510
: CITY: BOSTON
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Ascii (text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/465,590
: FILING DATE: 03-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/238,212
: FILING DATE: 02-MAY-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/121,438
: FILING DATE: 14-SEP-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/946,233
: FILING DATE: 14-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Myers, Paul L.
: REGISTRATION NUMBER: 35,695
: REFERENCE/DOCKET NUMBER: MPG-006C2DV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1386 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1386
US-08-465-590-3

Query Match 34.5%; Score 216.4; DB 3; Length 1386;
Best Local Similarity 67.9%; Pred. No. 1.9e-61;
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;
QY 146 ccagtggaaagatggaactgcatgtgtgtggattatccctgcatcagcttcaatgtcttaa 205
Db 173 CTACGGAAACAAAGTGTGATCTGTGGGATCATTTGTCATCGGGCCCAATGTGCTCA 232

CC enhancer function. An isolated clone was designated Ikaros and
 CC contained the 1788 bp sequence. A 300 bp 3'-terminal segment of
 CC this was used to identify 4 other Ikaros isoforms (see V42807-10).
 CC Different isoforms arise by differential splicing of Ikaros gene
 CC transcripts. Isoform mik-2 lacks exon 3. It is expressed in
 CC embryonic liver, thymus and brain, but only in thymus and spleen
 CC after birth. The Ikaros gene is located at the proximal arm of
 CC murine chromosome 11. Ikaros proteins are suggested to play a role
 CC as a genetic switch regulating entry into the T cell lineage. The
 CC murine and human (see V42806, V42811 and V42840) Ikaros sequences
 CC are very similar. The invention provides Ikaros nucleic acids,
 CC vectors and host cells expressing Ikaros proteins. These are used
 CC to treat T and B cell diseases (e.g. immune deficiencies caused by
 CC drugs, radiation or cancers), to control expression of heterologous
 CC genes placed under control of an Ikaros-responsive element, to
 CC treat nervous system diseases (e.g. Alzheimer's disease), and to
 CC modulate cell division, amplification or differentiation, especially
 CC in hematopoietic cells. Some Ikaros isoforms are antagonistic of
 CC others and may be used to inhibit interaction with DNA sequences.
 CC The same effect can be achieved with Ikaros-binding oligonucleotides.
 CC Examining the expression of the Ikaros gene, or its allelic
 CC structure, can be used to assess risk of acquiring the above
 CC diseases.
 SQ Sequence 1788 BP; 437 A; 488 C; 493 G; 370 T;

Query Match 27.1%; Score 170; DB 1; Length 1788;
 Best Local Similarity 67.1%; Pred. No. 3.8e-45;
 Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;
 QY 229 ggtgaagccattccagtgtaatacgtgtgggcatcttttactcagaagaagtaacctc 288
 DB 382 GGTGAAGCGCCTTCCAGTGCACACCAAGTCTGGGGCTCTCTTTACCCAGAAAGGCAACCTC 441
 QY 289 ctccgcccacataaactgcacacagcaggggaaacccctttaagtgtcacctctgcaactat 348
 DB 442 CTCCGGCACAATCAAGTGTGACTCGGCTGAGACGCCCTTCAATGCCATCTTTGCCAATAT 501
 QY 349 gcatgccaaagaagatgcgtcgcacagggggcatcttagcacacattctgtggagaaaccc 408
 DB 502 GCCTGCGCGGAGGAGCGCCCTCACCGGCCACCTGAGGACGCACCTCCGTTGTAAGCCT 561
 QY 409 tacaaatgtgaagttgtggaagagttacaacagcagaagaagttcccttgaggagcaag 468
 DB 562 CACAATGTGATATTGTGCCGAGCTATAACACGAGAGCTCTTTAGAGGAGCATAA 621
 QY 469 gagcgtgcggtacatttcttcagagcaactgacc-----caggggacactgcaagtgcg 522
 DB 622 GAGCGATGCCACAACACTACATTGGAAGCATGGGGCTTCCGGGGCTGTGCCCAGTCAATTAAG 681
 QY 523 gagggaagacacatcaag-----cagagatgggaagtgaagagct 564
 DB 682 GAAGAAACTAACCAACACAGAGATGGCAGAGAGACTGTGCAAGATAGGAGAGAGGTCC 741
 QY 565 ctctactgcacagattagcaagaatgtgcaaaacgaaaaagctcaatgctcagaag 624
 DB 742 CTGTCTGTCAGAGCTGGCGAAGCATGTGCCCCAAGCTAAGAGCTCTATGCTCTCAGAA 801
 QY 625 tt 626
 DB 802 TT 803

RESULT 15

ID T16061
 AC T16061; standard; CDNA; 1296 BP.
 DT 08-MAY-1996 (first entry)
 DE Murine Ikaros cDNA mik-3.
 KW Ikaros; transgene; transgenic animal; transgenic mouse; mik-3;
 KW immunocomprised; immune system disorder; nervous system disorder;
 KW animal model; ss.
 OS Mus musculus.

PN WO9604372-A1.
 PD 15-FEB-1996.
 PF 28-JUL-1995; U09345.
 PR 29-JUL-1994; US-283300.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 96-129389/13.
 DR P-PSDB; R92016.
 PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
 PT severely immuno-compromised and can be used as model to determine
 PT effects of treatment for immune and nervous system disorders
 PT effects of treatment for immune and nervous system disorders
 PS Disclosure: Page 63-64; 102pp; English.
 CC CDNA clones (T16059 and T16061-64) encode different isoforms,
 CC mik-2, mik-3, mik-1, mik-4 and mik-5 (R92014 and R92016-19,
 CC respectively), of the mouse Ikaros protein, a master regulator of
 CC hematopoietic differentiation. They were isolated from a mature
 CC T-cell line E-14 cDNA library. The isoforms arise by differential
 CC splicing of the Ikaros genomic locus. All include exons E1/2 and
 CC E7, but have different combinations of exons E3-E6 encoding a
 CC zinc finger domain. Transgenic animals, pref. mice, having a
 CC mutated Ikaros transgene, esp. a mutation that alters the DNA
 CC binding domain of the Ikaros protein, are used as models to
 CC determine the effects of treatments for immune or nervous system
 CC disorders.
 SQ Sequence 1296 BP; 321 A; 363 C; 365 G; 247 T;

Query Match 23.7%; Score 148.6; DB 1; Length 1296;
 Best Local Similarity 73.4%; Pred. No. 2.7e-38;
 Matches 190; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
 QY 146 ccagtggaagatgaactgcgatgtgtgattatcctgcacagcttcaatgtcttaa 205
 DB 338 CTACGGAAACTAAAGTGTGATATCTGTGGATCGTTTGATCGGCCCAATGTGCTCA 397
 QY 206 tgggtcataagcgaagccatactgggtgaacccattccagtgtaatacagtggtgggcat 265
 DB 398 TGGTTTCACAAAAGAGTATACACTGGTGAACGGCCTTTCCAGTCAACCACTCTGGGGCCT 457
 QY 266 cttttactcagaaggttaacctccctccgccacattaaactgcacacaggggaaaaacctt 325
 DB 458 CCTTTACCCAGAAAGGCAACCTCTCTCGGCACATCAAGCTGCACCTCGGTGAGAGCCCT 517
 QY 326 ttaagtgtcaacctgtgcaactatgcatgccaaagaagatgcgtctcacggggcatctta 385
 DB 518 TCAAAATGCCATCTTTGCAACTATGCTGCGCGGAGGAGCGCCCTCAACGCCACCTGA 577
 QY 386 ggacacattctgtggagaa 404
 DB 578 GGACGCACTCCGGAGACAA 596

Search completed: November 6, 1999, 02:38:45
 Job time: 18376 sec

QY 409 tacaattgtgagtttttgggaagagattacaagcagagagaagttcccttggagagacacaag 468
 DB 562 CACAATGTGATATTGTGGCGGAGCTATAAACAGCGAAGCTCTTTAGAGGAGCATATAA 621
 QY 469 gagcgtccgtacatttctcagagcactgacc-----caggggacactgcaagtgcg 522
 DB 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCGGCGGTGTGCCCAAGTCAATTAAG 681
 QY 523 gaggcaagacacatcaaaag-----cagagatggggaagtgaagagcgt 564
 DB 682 GAAGAACTAACCAACAGAGATGSCAGACAGCTGTCAAGATAGGAGAGAGGTCC 741
 QY 565 ctctactggcagattagcagaatgtgcaaaaacgaaagcctcaagaaa 624
 DB 742 CTCTGCTGGCAGGCTGCGCAAGCAATGCGCAACGCTATGAGAGCTCTATGCCTCAGAAA 801
 QY 625 tt 626
 DB 802 TT 803

RESULT 13
 V6968
 ID V6968 standard; cDNA; 1788 BP.
 AC V6968;
 DT 14-JAN-1999 (first entry)
 DE Murine Ikaros encoding cDNA mIk-2.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
 KW differentiation marker; immune system; corpus striatum; AIDS;
 KW Alzheimer's disease; ss.
 OS Mus sp.
 FH key Location/Qualifiers
 FT CDS 223..1518
 FT /tag= a
 FT /product= "mIk-2"
 FT /transl_except= (pos:385..387,aa:Gln)
 EN US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 463590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR P-PSDB; W72671.
 DR Ikaros poly:peptide(s) - useful for treating disorders of immune
 PT system or corpus striatum
 PS disclosure: Column 51-56; 11pp; English.
 CC The present invention describes a purified peptide having at least one
 CC of the following properties: (a) it stimulates transcription of a DNA
 CC sequence under the control of a delta A element, an NFkB element or an
 CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
 CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
 CC consensus sequence; (c) it competitively inhibits the binding of a
 CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
 CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
 CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
 CC (e) it inhibits protein-protein interactions of transcriptional complexes
 CC formed with naturally occurring Ikaros isoforms. The proteins, provided
 CC that they stimulate gene transcription under the control of delta A
 CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit binding of naturally occurring Ikaros isoforms to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
 CC inhibit protein-protein interactions of transcriptional complexes with
 CC naturally occurring Ikaros isoforms, can be used to treat immune system
 CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
 CC Alzheimer's disease. The present sequence encodes a specifically
 CC claimed mouse Ikaros protein.
 SQ Sequence 1788 BP; 437 A; 486 C; 495 G; 370 T;

Query Match 27.1%; Score 170; DB 1; Length 1788;
 Best Local Similarity 67.1%; Pred. No. 3.8e-45;
 Matches 283; Conservative 0; Mismatches 115; Indels 24; Gaps 2;
 QY 229 ggtgaagcccccattccagtgtaattcagtggtgggacatcttttactcagaagaagtaacctc 288
 DB 382 GGTGAAGGGCTTTCCAGTGCACACAGCTCTGGGCGCTCTTTACCCAGAAAGGCAACCTC 441
 QY 289 ctccgcccacattaaactgcacacaggggaaaaaccccttttaagtgtcacctctgcaactat 348
 DB 442 CTGGGACACATCAAGCTGCACCTCGGGTGAGAGCCCTTCAAAATGCCATCTTTGCAACTAT 501
 QY 349 gcatgccaaaagaagatgcgtcacggggcatcttaggcacacattctgtgggaaacccc 408
 DB 502 GCCTGCCCGCGGAGGAGCGCCCTCACGGCCACCTGAGGAGCAGCACCTCCGTGGTAAAGCCT 561
 QY 409 tacaattgtgagttttgtggaagagttacaagcagagaagttcccttggaggacacaag 468
 DB 562 CACAATGTGATATTGTGGCGGAGCTATAAACAGCGAAGCTCTTTAGAGGAGCATATAA 621
 QY 469 gagcgtccgtacatttctcagagcactgacc-----caggggacactgcaagtgcg 522
 DB 622 GAGCGATGCCACACTACTTGGAAAGCATGGCCCTTCCGGCGGTGTGCCCAAGTCAATTAAG 681
 QY 523 gagcgaagacacatcaaaag-----cagagatgggaagtgaagagcgt 564
 DB 682 GAAGAACTAACCAACAGAGATGSCAGACAGCTGTGCAAGATAGGAGAGAGGTCC 741
 QY 565 ctctactggcagattagcagaatgtgcaaaaacgaaagcctcaagtcctcagaaa 624
 DB 742 CTCTGCTGGCAGGCTGCGCAAGCAATGCGCAACGCTATGAGAGCTCTATGCCTCAGAAA 801
 QY 625 tt 626
 DB 802 TT 803

RESULT 14
 V42805
 ID V42805 standard; cDNA; 1788 BP.
 AC V42805;
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform mIk-2 cDNA.
 KW Ikaros; mIk-2; transcription factor; mouse; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Mus sp.
 FH key Location/Qualifiers
 FT CDS 223..1518
 FT /tag= a
 FT /transl_except= (pos:385..387, aa:Gln)
 EN CA2194256-A.
 PD 05-MAR-1998.
 PR 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR P-PSDB; W70963.
 PT New nucleic acid encoding Ikaros protein involved in early
 PT differentiation of lymphocytes - existing in several isoforms, and
 PT related products, used to treat e.g. immune diseases or cancer and
 PT to control cell differentiation
 PS Claim 7; page 68-70; 158pp; English.
 CC This is the nucleotide of mouse Ikaros cDNA (isoform mIk-2) that
 CC codes for a 431-amino acid zinc finger protein (see W70963) that is
 CC involved in the early differentiation of lymphocytes. A cDNA
 CC library constructed from the mature murine T cell line E14 was
 CC screened with a multimerised oligonucleotide (see V42829) derived
 CC from a protein binding site (see V42804) of the CD3-delta enhancer
 CC to identify T cell specific sequences that bind and mediate

CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC inhibit protein-protein interactions of transcriptional complexes with
CC naturally occurring Ikaros isoforms, can be used to treat immune system
CC disorders, e.g. leukemia or AIDS, or corpus striatum disorders, e.g.
CC Alzheimer's disease. The present sequence encodes a specifically
CC claimed mouse Ikaros protein.
SQ Sequence 2049 BP; 514 A; 531 C; 574 G; 430 T;

Query Match 33.4%; Score 209.8; DB 1; Length 2049;
Best Local Similarity 67.1%; Pred. No. 5.7e-58;
Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;
Qy 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagcttcaatgtcttaa 205
Db 560 CTAACGGAATAAAGTGTGATATCTCTGGGATCGTTTGGCAGGCGCCCAATGTGCTCA 619
Qy 206 tgggttcataagcgaagccatacgtgtgaacccattccagtgtaatacagtggtggcat 265
Db 620 TGGTTCACAAAGAAGTGTGATGTAACCGCTTCCAGTGCACAGTCTGGGGCCT 679
Qy 266 cttttactcagaagaagtaacctctccgccacattaaactgcacacaggggaaacacctt 325
Db 680 CTTTACCAGAAAGCAACCTCTCGCGCACATCAACCTGCACCTCGGGTGAAGCCCT 739
Qy 326 ttaagtgtaacctctgcactatgcataatgcacaaagaagatgcgtcaccggggcatctta 385
Db 740 TCAATGCCATCTTTGCAACTATGCTGCGCGGAGGAGCGCCCTACCGCGCACCTGA 799
Qy 386 ggacacattctgtgagaacccataatgtgattgtttgtggaaggttacaagcaga 445
Db 800 GGACGCATCTCGTGGTAAGCTTCAAAATGTGGATATTGTGGCGGAGCTATAAACAGC 859
Qy 446 gaagttcccttgaggagcacaagggagcgctgcgtacatttctcagagcactgacc--- 503
Db 860 GAAGCTCTTTAGAGAGAGATAAAGAGCCATGCCACAACTACTTGCAGAGCATGGCCCTTC 919
Qy 503 ---caggggacactgcaagtgcggaggcgaacacacatcaag----- 542
Db 920 CGGGCGTGTGCCAGTCAATTAAGGAAGAACTAACCAACAGAGATGGCAGAACCTGT 979
Qy 542 cagagatgggaagtgaagaactctcgtactggacagatgacagatgaagcaatgtgcaaac 601
Db 980 GCAAGATAGGAGCAGAGAGTGCCTTGTCTGTGGACAGGCTGGCAGCAATGTGCGCAAC 1039
Qy 602 gaaaagctcaatgcctcagaatt 626
Db 1040 GTAAGAGCTCTATGCTCAGAAATT 1064

RESULT 10
V42808
ID V42808 standard; cDNA; 2049 BP.
AC V42808;
DT 11-JAN-1999 (first entry)
DE Mouse Ikaros isoform mik-1 cDNA.
KW Ikaros; mik-1; transcription factor; mouse; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis; ss.
OS Mus sp.
FR Key Location/Qualifiers
FT CDS 223..1779
FT exon /*tag= a
FT exon 223..384
FT exon /*tag= b
FT exon /*tag= c
FT exon 385..643
FT exon /*tag= Ex3
FT exon 644..810
FT exon /*tag= d
FT exon /*tag= Ex4
FT exon 811..933

FT /*tag= e
FT /number= Ex5
FT 934..1076
FT /*tag= f
FT /number= Ex6
FT 1077..1779
FT /*tag= g
FT /number= Ex7
PN CA2194256-A.
PD 05-MAR-1998.
PF 02-JAN-1997; 194256.
PR 05-SEP-1996; US-711417.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR P-PSDB; W70966.
PT New nucleic acid encoding Ikaros protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer, and
PT to control cell differentiation
PS Claim 7; Page 75-77; 158pp; English.
CC This is the nucleotide of mouse Ikaros cDNA (isoform mik-1) that
CC codes for a 518-amino acid zinc finger protein (see W70966) that is
CC involved in the early differentiation of lymphocytes. mik-1 cDNA
CC was isolated from a mature murine T cell line E14 library using a
CC 300 bp fragment from the 3' end of mik-2 cDNA (see V42805) as
CC probe. 5 Different isoforms of mouse Ikaros (see V42805 and
CC V42807-10) have been identified. These arise by differential
CC splicing of Ikaros gene transcripts. Isoform mik-1 contains all 7
CC exons. It is abundantly expressed in the early foetal liver, the
CC maturing thymus and the postnatal spleen. The Ikaros gene is
CC located at the proximal arm of murine chromosome 11. Ikaros
CC proteins are suggested to play a role as a genetic switch
CC (see V42806, V42811 and V42840) Ikaros sequences are very similar.
CC The invention provides Ikaros nucleic acids, vectors and host cells
CC expressing Ikaros proteins. These are used to treat T and B cell
CC diseases (e.g. immune deficiencies caused by drugs, radiation or
CC cancers), to control expression of heterologous genes placed under
CC control of an Ikaros-responsive element, to treat nervous system
CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
CC amplification or differentiation, especially in haematopoietic
CC cells. Some Ikaros isoforms are antagonistic of others and may be
CC used to inhibit interaction with DNA sequences. The same effect
CC can be achieved with Ikaros-binding oligonucleotides. Examining
CC the expression of the Ikaros gene, or its allelic structure, can be
CC used to assess risk of acquiring the above diseases.
SQ Sequence 2049 BP; 514 A; 531 C; 574 G; 430 T;

Query Match 33.4%; Score 209.8; DB 1; Length 2049;
Best Local Similarity 67.1%; Pred. No. 5.7e-58;
Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;
Qy 146 ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagcttcaatgtcttaa 205
Db 560 CTAACGGAATAAAGTGTGATATCTCTGGGATCGTTTGGCAGGCGCCCAATGTGCTCA 619
Qy 206 tgggttcataagcgaagccatacgtgtgaacccattccagtgtaatacagtggtggcat 265
Db 620 TGGTTCACAAAGAAGTGTGATGTAACCGCTTCCAGTGCACAGTCTGGGGCCT 679
Qy 266 cttttactcagaagaagtaacctctccgccacattaaactgcacacaggggaaacacctt 325
Db 680 CTTTACCAGAAAGCAACCTCTCGCGCACATCAACCTGCACCTCGGGTGAAGCCCT 739
Qy 326 ttaagtgtaacctctgcactatgcataatgcacaaagaagatgcgtcaccggggcatctta 385
Db 740 TCAATGCCATCTTTGCAACTATGCTGCGCGGAGGAGCGCCCTACCGCGCACCTGA 799
Qy 386 ggacacattctgtgagaacccataatgtgattgtttgtggaaggttacaagcaga 445
Db 800 GGACGCATCTCGTGGTAAGCTTCAAAATGTGGATATTGTGGCGGAGCTATAAACAGC 859

Db 818 AACGTAAGAGCTCTATGCTCAGAAATT 845

RESULT 8

T16062
ID T16062 standard; cDNA; 2049 BP.
DT T16062:
DE Murine Ikaros cDNA mik-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; mik-1;
KW immunocompromised; immune system disorder; nervous system disorder;
KW animal model; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 223..1779
FT /*tag= a
FT exon 223..381
FT /*tag= b
FT /*label= Exon-1/2
FT exon 382..642
FT /*tag= c
FT /*label= Exon-3
FT exon 643..810
FT /*tag= d
FT /*label= Exon-4
FT exon 811..933
FT /*tag= e
FT /*label= Exon-5
FT exon 934..1068
FT /*tag= f
FT /*label= Exon-6
FT exon 1069..1776
FT /*tag= g
FT /*label= Exon-7

WO9604372-A1.

15-FEB-1996.

28-JUL-1995. U09345.

29-JUL-1994; US-283300.

(GEO) GEN HOSPITAL CORP.

Georgopoulos K;

WPI; 96-129389/13.

P-PSDB; R92017.

Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immunocompromised and can be used as model to determine effects of treatment for immune and nervous system disorders

Disclosure: Page 65-67; 102pp; English.

C DNA clones (T16059 and T16061-64) encode different isoforms, mik-2, mik-3, mik-1, mik-4 and mik-5 (R92014 and R92016-19, respectively), of the mouse Ikaros protein, a master regulator of hematopoietic differentiation. They were isolated from a mature T-cell line E-14 cDNA library. The isoforms arise by differential

splicing of the Ikaros genomic locus. All include exons E1/2 and E7, but have different combinations of exons E3-E6 encoding a zinc finger domain. Transgenic animals, pref. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to

determine the effects of treatments for immune or nervous system disorders.

Sequence 2049 BP; 514 A; 531 C; 574 G; 430 T;

Query Match 33.4%; Score 209.8; DB 1; Length 2049;

Best Local Similarity 67.1%; Pred. No. 5.7e-58;

Matches 339; Conservative 0; Mismatches 142; Indels 24; Gaps 2;

QY 146 ccagtggaagatgaactcgcgtgtgtgtgattatccctgcacagcttcaatgcttaa 205

Db 560 CTAACGGAATAACTAAGTGTGATATCTGTGGATCGTTTGATCGGCCCAATGTGCTCA 619

QY 206 tgggtcataagcgaagccactactgtgtggaagccattccagtgtaatcagtggtgggcat 265

Db 620 TGGTTCACAAAGAAGTACATCTGGTGAACGGCTTTCCAGTGCAACAGTCTGGGCT 679

QY 266 cttttactcagaagtgtaacctctcctccacattaaactgcacacaggggaaacctt 325
Db 680 CCTTTACCCAGAAAGGCAACCTCTCGCGCACATCAAGCTGCACCTCGGAGAGCCCT 739
QY 326 ttaagtgcacctctgaactatgcattcacaagaagagatgcgtcaccggggcatetta 385
Db 740 TCAATGCCATCTTTGCAACTATGCTCGCGGAGGAGCGCCTCACCGCCACCTGA 799
QY 386 ggacacattctgtggagaaacccacaaatgtgagttttgtggaagagattacaagcaga 445
Db 800 GGACCCACTCCGTTGGTAAGCCTCACAATCTGGATATTTGGCCGGAGACTATAACAGC 859
QY 446 gaagttcccttgaggagacacaagagcgctgcctacattcttcagagcactgacc--- 503
Db 860 GAAGCTCTTTAGAGGAGCATAAAGAGCGCATGCCACAACACTACTTGGAAAGCATGGCCCTTC 919
QY 503 ---caggggacactgcaagtgcggaggcaagacacatcaaaag----- 542
Db 920 CGGCGTGTGCCAGTCATTAAAGGAAGAACTAACCAACAGAGATGCGCAGAAGACCTGT 979
QY 542 cagagatgggaagtgaagagctctcgtactggacagagattagcaagcaatgtggcaaac 601
Db 980 GCAAGATAGGAGCAGAGAGGTCCTTGTCTTGGACAGGCTGCGAAGCAATGTCCGCCAATC 1039
QY 602 gaaaaagctcaatgcctcagaaatt 626
Db 1040 GTAGAGCTCTATGCTCAGAAATT 1064

RESULT 9

V66971

ID V66971 standard; cDNA; 2049 BP.

AC V66971;

DE 14-JAN-1999 (first entry)

Mouse Ikaros encoding cDNA mik-1.

CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;

differentiation marker; immune system; corpus striatum; AIDS;

Alzheimer's disease; ss.

OS Mus sp.

FH Key Location/Qualifiers

FT CDS 223..1779

FT /*tag= a

FT /product= "mik-1"

US5824770-A.

PD 20-OCT-1998.

PF 05-JUN-1995; 465590.

PR 02-MAY-1994; US-238212.

PR 14-SEP-1992; US-946233.

PR 14-SEP-1993; US-121438.

PR 05-JUN-1995; US-465590.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K;

WPI; 98-582621/49.

P-PSDB; W72674.

Ikaros poly-peptide(s) - useful for treating disorders of immune system or corpus striatum

The present invention describes a purified peptide having at least one of the following properties: (a) it stimulates transcription of a DNA

sequence under the control of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (c) it competitively inhibits the binding of a

naturally occurring Ikaros isoform to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (d) it

competitively inhibits Ikaros binding to Ikaros responsive elements; or (e) it inhibits protein-protein interactions of transcriptional complexes

formed with naturally occurring Ikaros isoforms. The proteins, provided that they stimulate gene transcription under the control of delta A

elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

competitively inhibit binding of naturally occurring Ikaros isoforms to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

Best Local Similarity 67.9%; Pred. No. 3.5e-60;		Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;	
Qy	146	ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagcttcaatgtcttaa	205
Db	173	CTACGGAAACTAAAGTGTGATATCTGTGGATCAITTTGGATCGGCCCAATGTGCTCA	232
Qy	206	tgggttaagcgaagcactactgtgtgaagccattccagtgtaactgaagtggtggcat	265
Db	233	TGGTTCAAAAGAAAGCCACACTGGAGAACGGCCCTTCCAGTGAATCAGTGGCGGCCCT	292
Qy	266	cttttactcagaaggttaacctctcgcgcacattaaactgcacacaggggaaaaacctt	325
Db	293	CATTACCCAGAGGGCAACCTGTCTCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCT	352
Qy	326	ttaagtgctacctctgcaactatcatgccaagaagagatgcgctcacggtgcatctta	385
Db	353	TCAATGCCACCTCTGCAACTACGCTGCGCGGAGGGAGCCCTCACTGGCCACCTGA	412
Qy	386	ggacacattctgtggagaaacccctacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt	445
Db	413	GGAGCGCACTCGTGTGTAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAACAGC	472
Qy	446	gaagttcccttgaggacacagagcgcctgcctacattttcttcagagcactgacc--c	503
h	473	GAAGCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTTGGAAAGCATGGCCCTT	532
Qy	504	aggggacactgcaagtgcggaggcaagacacatcaaaagca-----	544
Db	533	CGGCGACACTGTACCCAGTCAAAAAGAAAGAACTAAGCAGTGAATGCCAGAAGACC	592
Qy	544	-----gagatgggaagtgaagagactctgtactggacagattagcaagaatgtggcaa	598
Db	593	TGTCAAGATAGGATCAGAGAGATCTCTCGTGTGGACAGACTAGCAAGTAATGTGCCCA	652
Qy	599	aacgaaaaagctcaatgcctcagaatt 626	
Db	653	AACGTAGAGCTCTATGCTCAGAAAT 680	
RESULT 7			
ID	V42840	standard; cDNA; 1551 BP.	
AC	V42840	11-JAN-1999 (first entry)	
DE	Human Ikaros isoform hlk-1 cDNA.		
KW	Ikaros; hlk-1; transcription factor; human; lymphocyte;		
KW	cell differentiation; T cell; cancer; immunodeficiency;		
KW	Alzheimer's disease; therapy; diagnosis; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	exon	1..26	
FT	exon	/*tag= a	
FT	exon	/number= Ex1	
FT	exon	27..163	
FT	exon	/*tag= b	
FT	exon	/number= Ex2	
FT	exon	164..420	
FT	exon	/*tag= c	
FT	exon	/number= Ex3	
FT	exon	421..588	
FT	exon	/*tag= d	
FT	exon	/number= Ex4	
FT	exon	589..714	
FT	exon	/*tag= e	
FT	exon	/number= Ex5	
FT	exon	715..849	
FT	exon	/*tag= d	
FT	exon	/number= Ex6	
FT	exon	850..1551	
FT	exon	/*tag= e	
FT	exon	/number= Ex7	
PN	CA2194256-A.		

PD	05-MAR-1998.	194256.	
PF	02-JAN-1997.	US-711417.	
PA	(GEHO) GEN HOSPITAL CORP.		
PI	Georgopoulos K;		
DR	WPI; 98-378292/33.		
DR	P-PSDB; W70971.		
PT	New nucleic acid encoding Ikaros protein involved in early differentiation of lymphocytes - existing in several isoforms, and related products, used to treat e.g. immune diseases or cancer and to control cell differentiation		
PS	Claim 1; Page 127-129; 158pp; English.		
CC	This is the nucleotide of human Ikaros cDNA (isoform hlk-1) that codes for a 516-amino acid zinc finger protein (see W70971) that is involved in the early differentiation of lymphocytes. It was isolated from a Jurkat T cell line library using mouse Ikaros exon 7 cDNA as probe. The Ikaros gene maps between p11.2-p13 on human chromosome 7. The human and murine Ikaros sequences (see V42805-11 and V42840) are highly conserved. Differential splicing of Ikaros gene transcripts gives rise to different Ikaros protein isoforms. The invention provides Ikaros nucleic acids, vectors and host cells expressing Ikaros proteins. These are used to treat T and B cell diseases (e.g. immune deficiencies caused by drugs, radiation or cancers), to control expression of heterologous genes placed under control of an Ikaros-responsive element, to treat nervous system diseases (e.g. Alzheimer's disease) and to modulate cell division, amplification or differentiation, especially in haematopoietic cells. Some Ikaros isoforms are antagonistic of others and may be used to inhibit interaction with DNA sequences. The same effect can be achieved with Ikaros-binding oligonucleotides. Examining the expression of the Ikaros gene, or its allelic structure, can be used to assess risk of acquiring the above diseases.		
SQ	Sequence 1551 BP; 393 A; 450 C; 265 T;		
Query Match 34.5%; Score 216.4; DB 1; Length 1551;			
Best Local Similarity 67.9%; Pred. No. 3.7e-60;			
Matches 345; Conservative 0; Mismatches 136; Indels 27; Gaps 2;			
Qy	146	ccagtggaagatgaactgcgatgtgtgtgattatctctgcatacagcttcaatgtcttaa	205
Db	338	CTAAGGAAACTAAAGTGTGATATCTGTGGATCATTTGATCGGCCCAATGTGCTCA	397
Qy	206	tgggttaagcgaagcactactgtgtgaagccattccagtgtaactgaagtggtggcat	265
Db	398	TGGTTCAAAAGAAAGCCACACTGGAGAACGGCCCTTCCAGTGAATCAGTGGCGGCCCT	457
Qy	266	cttttactcagaaggttaacctctcgcgcacattaaactgcacacaggggaaaaacctt	325
Db	458	CATTACCCAGAGGGCAACCTGTCTCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCT	517
Qy	326	ttaagtgtaacctctgcaactatgcatgccaagaagagatgcgtcacgggcatctta	385
Db	518	TCAATGCCACCTCTGCAACTACGCTGCGCGGAGGGAGCCCTCACTGGCCACCTGA	577
Qy	386	ggacacattctgtggagaaacctacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt	445
Db	578	GGAGCGCACTCCGTTGTTAAACCTCACAAATGTGGATATTGTGCCGAAGCTATAACAGC	637
Qy	446	gaagttcccttgaggagacacagagcgtgcggtacattttcttcagagcactgacc--c	503
Db	638	GAACGCTCTTTAGAGGAACATATAAGAGCGCTGCCACAACTACTTTGGAAAGCATGGCCCTT	697
Qy	504	aggggacactgcaagtgcggaggaagacacatcaaaagca-----	544
Db	698	CGGCGACACTGTACCCAGTCAATTAAGAGAAACTAAGCAGTGAATGCCAGAAGACC	757
Qy	544	-----gagatgggaagtgaagagctctctgtactggacagattagcaagaatgtggcaa	598
Db	758	TGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACAGACTAGCAAGTAATGTGCCCA	817
Qy	599	aacgaaaaagctcaatgcctcagaatt 626	

FT CDS 1. .1386
 FT /*tag= a
 FT /product= "Ikaros"
 PD US5824770-A.
 PD 20-OCT-1998. 465590.
 PR 05-JUN-1995; US-238212.
 PR 02-MAY-1994; US-946233.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-582621/49.
 DR P-PSDB: W72672.
 PT Ikaros poly:peptide(s) - useful for treating disorders of immune
 PT system or corpus striatum
 PS Disclosure; Column 55-58; ilipp; English.
 CC The present invention describes a purified peptide having at least one
 CC of the following properties: (a) it stimulates transcription of a DNA
 CC sequence under the control of a delta A element, an NFkB element or an
 CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
 CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
 CC consensus sequence; (c) it competitively inhibits the binding of a
 CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
 CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
 CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
 CC (e) it inhibits protein-protein interactions of transcriptional complexes
 CC formed with naturally occurring Ikaros isoforms. The proteins, provided
 CC that they stimulate gene transcription under the control of delta A
 CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit binding of naturally occurring Ikaros isoforms to
 CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
 CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
 CC inhibit protein-protein interactions of transcriptional complexes with
 CC naturally occurring Ikaros isoforms, can be used to treat immune system
 CC disorders, e.g. leukemia or AIDS, or corpus striatum disorders, e.g.
 CC Alzheimer's disease. The present sequence encodes a specifically
 CC claimed human Ikaros protein.
 SQ Sequence 1386 BP; 350 A; 403 C; 395 G; 238 T;

Query Match 34.5%; Score 216.4; DB 1; Length 1386;
 Best Local Similarity 67.9%; Pred. No. 3.5e-60;
 Matches 34; Conservative 0; Mismatches 136; Indels 27; Gaps 2;

Qy 146 ccagtgaagatgaactgcgtgtgtgattctctgcagtcagcttcaatgtcttaa 205
 Db 173 CTAACGGAATACTAAGTGTATCTGTGGATCATTTGTCATCGGGCCCAATGTCTCA 232
 Qy 206 tgggtcataagcgaagccatactggtgaaocccattccagtgtaactagtggtgggcat 265
 Db 233 TGGTTTCAAAAGAAAGCCACACTGAGAACGGCCCTTCCAGTGCATCAGTGGCGGCT 292
 Qy 266 cttttactcagaaggaagtaacctctccgcacattaaactcacacaggggaaacatt 325
 Db 293 CATTACCCAGGAGGAACTGCTCGGGCACATCAAGCTGCATTCGCGGAGAGGCCCT 352
 Qy 326 ttaagtgtcactctgcaactatgatcgccaaagaagatgcgtcacggggcatctta 385
 Db 353 TCAATGCCACCTCTGCAACTACGCCCTGCCCGGAGGGAGGCCCTCAGTGGCCACCTGA 412
 Qy 386 ggcacattctgtggagaaacccataactgtgagttttgtggagaggtttacagcaga 445
 Db 413 GGAGCGCACTCGTGGTAAAGCCTCACAATGTGGATATTGTGGCGGAAGCTATAAAGAGC 472
 Qy 446 gaagttcccttgagagcagaagagcgtccgtacattcttcttcagagcactacc-c 503
 Db 473 GAACGCTTTTAGAGAACTAAGAGGCGCTGCCACAACTACTTGGAAAGCATGGCGCTTC 532
 Qy 504 aggggacactgcaagtgcgggggaagacacacatcaaaagca----- 544
 Db 533 CGGGCACACTGTACCCAGTCTAATTAAGAGAACTAAGACAGTCAATGGCAGAGACC 592

Qy 544 -----gagatgggaagtgaagagctctctactgacagattagcaagaatgtggcaa 598
 Db 593 TGTGCAAGATAGGATCAGAGACTCTCTGCTGTGCAGACACTAGCAAGTAAATGTGCCCA 652
 Qy 599 aacgaaaaagcctaactgcctcagaatt 626
 Db 653 ACGTAAGAGACTCTATGCTCTCAGAAAT 680
 RESULT 6
 V42806
 ID V42806 standard; cDNA; 1386 BP.
 AC V42806;
 DT 11-JAN-1999 (first entry)
 DE Human Ikaros isoform h1k-1 cDNA.
 KW Ikaros; h1k-1; transcription factor; human; lymphocyte;
 KW cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 1. .255
 FT /*tag= a
 FT /number= Ex3
 FT FT 256. .423
 FT /*tag= b
 FT FT /number= Ex4
 FT FT 424. .549
 FT /*tag= c
 FT FT /number= Ex5
 FT FT 550. .684
 FT /*tag= d
 FT FT /number= Ex6
 FT FT 685. .1386
 FT /*tag= e
 FT FT /number= Ex7
 PN CA2194256-A.
 PD 05-MAR-1998.
 PF 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 98-378292/33.
 DR P-PSDB: W70964.
 PT New nucleic acid encoding Ikaros protein involved in early
 PT differentiation of lymphocytes - existing in several isoforms, and
 PT related products, used to treat e.g. immune diseases or cancer and
 PT to control cell differentiation
 PS Claim 7: Page 70-72: 158pp; English.
 CC This is the nucleotide of human Ikaros cDNA (isoform h1k-1) that
 CC codes for a 461-amino acid zinc finger protein (see W70964) that is
 CC involved in the early differentiation of lymphocytes. It was
 CC isolated from a Jurkat T cell line library using mouse Ikaros exon
 CC 7 cDNA as probe. The Ikaros gene maps between p11.2-pl3 on human
 CC chromosome 7. The human and murine Ikaros sequences (see V42805-11
 CC and V42840) are highly conserved. Differential splicing of Ikaros
 CC gene transcripts gives rise to different Ikaros protein isoforms.
 CC The invention provides Ikaros nucleic acids, vectors and host cells
 CC expressing Ikaros proteins. These are used to treat T and B cell
 CC diseases (e.g. immune deficiencies caused by drugs, radiation or
 CC cancers), to control expression of heterologous genes placed under
 CC control of an Ikaros-responsive element, to treat nervous system
 CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
 CC amplification or differentiation, especially in haematopoietic
 CC cells. Some Ikaros isoforms are antagonistic of others and may be
 CC used to inhibit interaction with DNA sequences. The same effect
 CC can be achieved with Ikaros-binding oligonucleotides. Examining
 CC the expression of the Ikaros gene, or its allelic structure, can be
 CC used to assess risk of acquiring the above diseases.
 SQ Sequence 1386 BP; 352 A; 403 C; 397 G; 234 T;

Query Match

34.5%; Score 216.4; DB 1; Length 1386;

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	628	100.0	628	1	T60491	Human Aiolos parti
2	493.6	78.6	1984	1	T60490	Mouse Aiolos cDNA
3	216.4	34.5	1611	1	Q4980	Human Ikaros pepti
4	216.4	34.5	1386	1	T15060	Human Ikaros cDNA
5	216.4	34.5	1386	1	V65969	Human Ikaros encod
6	216.4	34.5	1386	1	V42806	Human Ikaros isofo
7	216.4	34.5	1551	1	V42840	Human Ikaros isofo
8	209.8	33.4	2049	1	T15062	Murine Ikaros cDNA
9	209.8	33.4	2049	1	V65971	Mouse Ikaros encod
10	209.8	33.4	2049	1	V42808	Mouse Ikaros isofo
11	170	27.1	1788	1	Q49579	Murine Ikaros gene
12	170	27.1	1788	1	T16059	Murine Ikaros cDNA
13	170	27.1	1788	1	V65968	Murine Ikaros cDNA
14	170	27.1	1788	1	V42805	Mouse Ikaros isofo
15	148.6	23.7	1296	1	T16061	Murine Ikaros cDNA
16	148.6	23.7	1296	1	V65970	Mouse Ikaros encod
17	148.6	23.7	1296	1	V42807	Mouse Ikaros isofo
18	107.6	17.1	1170	1	T16063	Murine Ikaros cDNA
19	107.6	17.1	1170	1	V65972	Mouse Ikaros encod
20	107.6	17.1	1170	1	V42809	Mouse Ikaros isofo
21	106.6	17.0	1004	1	T16065	Ikaros cDNA. Trans
22	106.6	17.0	1004	1	V65974	Ikaros isoform enc
23	106.6	17.0	1004	1	V42811	Human Ikaros cDNA
24	106.2	16.9	1168	1	V67125	Ikaros protein enc
25	104.6	16.7	168	1	V67126	Ikaros protein enc
26	71.6	11.4	1663	1	X39662	Renal cancer assoc
27	68.4	10.9	323	1	T23378	Human gene signatu
28	66.4	10.6	765	1	X39735	Gastric cancer ass
29	64.6	10.3	851	1	Q00648	Human secreted pro
30	63.8	10.2	345	1	T25341	Human gene signatu
31	63.8	10.2	353	1	V86582	EST clone AW95. Ne
32	62.8	10.0	398	1	O59865	Human brain Expres
33	62.8	10.0	1683	1	V73484	Human SRE-ZBP anal
34	62.8	10.0	2168	1	V73485	Human SRE-ZBP anal
35	62	9.9	2133	1	V01890	Human OTK18 gene.
36	62	9.9	3754	1	V01891	Human OTK18 gene.
37	61.6	9.8	264	1	V83634	DNA encoding a nuc
38	61.6	9.8	264	1	V98467	DNA encoding an an
39	61.6	9.8	264	1	X16976	Zinc finger protei
40	60.2	9.6	270	1	V89789	EST clone CO1069.
41	59.8	9.5	2680	1	V64579	Myc-binding zinc-f
42	59.6	9.5	135	1	V67122	Ikaros protein enc
43	58.2	9.3	3099	1	V64361	Human stem cell zi

56215 AATTACCCCTGTCCCGGAGAGAGCCCTGACCGGACATCTACGCACGCACGCAG 56269

Search completed: November 6, 1999, 06:39:51
Job time: 4705 sec


```
VERSION      U092199.1  GI:2062737
KEYWORDS
SOURCE       rainbow trout.
ORGANISM     Oncorhynchus mykiss
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
              Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
              Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE    1 (bases 1 to 2079)
AUTHORS      Hansen,J.D.
TITLE        Isolation and characterization of Ikaros homologues in the rainbow
              trout
JOURNAL      Unpublished
AUTHORS      Hansen,J.D.
TITLE        Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
              Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
FEATURES     Location/Qualifiers
              1..2079
               /organism="Oncorhynchus mykiss"
               /strain="shasta"
               /db_xref="taxon:8022"
               /cell_type="thymocyte"
              81..1232
               /gene="ikaros"
              81..1232
               /gene="ikaros"
               /note="ik-8 isoform; similar to mouse and human
              ikaros/lyf-1; alternatively spliced form missing exons 3
              and 6"
               /codon_start=1
               /product="ikaros homolog"
               /protein_id="AAB53433.1"
               /db_xref="PID:g2062738"
               /db_xref="GI:2062738"
               /translation="MEMEAEOMSPGDRGPPPPNDVSEENDEAMPIDLSASSNLO
              HNNRGDKELGEPFOCTQCGASTQKGNLLRHKLHSGKPFKCHLONACRRDAL
              SCHLTHSVGPKHCAYGKRSYKQSSLEEKHERCHNYLQCMGLONSIYTEKRFNSI
              SFEGPGELMOPHVIDQAINLGAESURPLIOTSPSTDGMVGMSMYPLHKPPA
              EGHGUSAKDSALLLLAKSASERKSGDSQSDSTDSNNEKAGVAGSLI
              YLTNHTDGVYVLPVKEQOQYEMRASIEIASGFKVLSGEGQVRAVRECHC
              RILFLDHVYTHMGCHGFROPFCNLCGRSQRDFESSHMTGRHRY"
BASE COUNT   593 a 500 c 464 g 521 t
ORIGIN
Query Match      25.3%; Score 158.6; DB 4; Length 2079;
Best Local Similarity 69.6%; Pred. No. 2e-35;
Matches 215; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 229 ggtgaagccattccagtgtaatacagtggtgggcatctttactcagaagaagtaacctc 288
      || || || || || || || || || || || || || || || || || || || || ||
DB 243 GGAGACGTCATTCAGTGCACCCAGTGTGGCGCTTCTTACCAGAGAAAGCAACCTG 302
      || || || || || || || || || || || || || || || || || || || || ||
QY 289 ctccgcacataaactcacacagaggggaaacacctttaagtgtcacctctgcaactat 348
      || || || || || || || || || || || || || || || || || || || || ||
DB 303 CTACGTACATCAATCAAGCTCCACTCAGAGAGAACCCCTCAAGGTGCACCTGTGCAACTAT 362
      || || || || || || || || || || || || || || || || || || || || ||
QY 349 gcatgcgaagaagagatgcgtccacggggcatcttaggacacatctgtggagaacccc 408
      || || || || || || || || || || || || || || || || || || || || ||
DB 363 GCCTGTGCGGGAGACACGCCCTCAGCGGTACCTGCGTACCCACTCTGTGGAAACCC 422
      || || || || || || || || || || || || || || || || || || || || ||
QY 409 tacaatgtgagtttggaggaggtttacaagcagagaagttcccttgaggagcacaag 468
      || || || || || || || || || || || || || || || || || || || || ||
DB 423 CACAAGTGTGCTTACTGTGGCGGAGCTACAGCAGCGTAGCTCTCTGGAGAACACAAG 482
      || || || || || || || || || || || || || || || || || || || || ||
QY 469 gacgcgtgcctacattttcagacactgaccacgggggacactcaagtgcgaggaca 528
      || || || || || || || || || || || || || || || || || || || || ||
DB 483 GAGCGGTGTGCACAACTACCTTCCAGTGCATGGGCTCCAGAACAGCATCTATACAGTGTAG 542
      || || || || || || || || || || || || || || || || || || || || ||
QY 529 agacacatc 537
      || || || ||
DB 543 AAAGGCTTC 551
      || || || ||
```

```

CDS
<1..>415
/genes="Ikaros"
/notes="similar to mouse and human Ikaros/LyF-1; composed
of exons 4, 5 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53435.1"
/db_xref="PID:g2062744"
/db_xref="GI:2062744"
/translation="GERPFCNOCGASFTQKGNLLRHILHSGEKFCHKMCMYACRR
RDALSGHLRTHSVGKPKHKGYSYKQSSLEHKECHNYLQSIGLQGLHLYAVKEE
SKTDAEDLSKIGSERSLVDLRASNAVKRKSMPQ"
BASE COUNT 124 a 93 c 100 g 98 t
ORIGIN

Query Match 26.5%; Score 166.2; DB 4; Length 415;
Best Local Similarity 65.3%; Pred. No. 1.4e-37;
Matches 27; Conservative 0; Mismatches 123; Indels 21; Gaps 1;

Qj 229 ggtgaacgcccattccagtgtaacagtggtgggcatcttttactcagaaggttaacctc 288
Nh 1 GGTGAACGGCCTTTCCAGTGCACACAGTGTGGTGCCTCTTTACCCAGAGGGCAATCTC 60

Qy 289 ctccgcccattaaactgcacacaggggaaaaaaccttttaagtgcaccttgcaactat 348
Db 61 CTTCTCATATCAAAATACACTCTGGTGAAAGGCCCTTTAAATGTCACATGTGTAACAT 120

Qy 349 gcatgccaagaagagatgcgtccacgggcatcttaggcacacattctgtggagaacccc 408
Db 121 GCATGCGAGCGCGAGGATGCATCTAGTGGTCACCTCAGGACACATTTCTAGGCAACCT 180

Qy 409 tacaatgtgagtttgggaaggagtacaagcagagaagttcccttgaggagcacaaag 468
Db 181 CACAAGTGTGATACTGTGGTCGTAGCTATAAGCAACGGAGCTCCTTAGAAGAGCATAAA 240

Qy 469 gaggcgtgcgtacattcttcagagacactgacccaggggacactgcaagtgcggagcca 528
Db 241 GAGAGATGTCAAACTACCTACAAAGTATAGGCTTGCGAGGCCCATCTTTATGCTGTAAA 300

Qy 529 agacacatcaaaa-----gcagagatgggaagtgaagagctctc 567
Db 301 GAAGAGACAAATATATATGGCAGAGACCTGTCCAGATGGGTGAGAGATCACTT 360

Qy 568 gtactgcagagattagcaagcaatgtggcaaaacgaaaagctcaatgcctcaga 622
Db 361 GTGCTTGACAGACTGTCGAAGTAAGCTTGCCAAAGTGAAGAGCTCTATGCCTCAGA 415

RESULT 13
OMU92200
LOCUS OMU92200 2183 bp mRNA VRT 02-MAY-1997
DEFINITION Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, Ik-2 isoform,
complete cds.
ACCESSION U92200
NID g2062739
VERSION U92200.1 GI:2062739
KEYWORDS rainbow trout.
SOURCE Oncorhynchus mykiss
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
Hansen, J.D.
Isolation and characterization of Ikaros homologues in the rainbow
trout.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2183)
AUTHORS Hansen, J.D.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland

```

```

FEATURES
Source
Location/Qualifiers
1..2183
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
50..1336
/gene="Ikaros"
50..1336
/gene="Ikaros"
/notes="Ik-2 isoform: similar to mouse and human
Ikaros/LyF-1; alternatively spliced form missing exon 3"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53434.1"
/db_xref="PID:g2062740"
/db_xref="GI:2062740"
/translation="MEMEPAEWSQMPGRDSPNDVSEENDEAMPIDELSSNLQ
HNRRGKELGERPFOCTCGASFTQKGNLLRHILHSGEKFCHKLCNYACRRRAL
SGHLRTHSVGKPKHCAICGRSYKQSSLEHKECHNYLOCMLQNSIYTVKEESNQ
NEQREDLSQMSKRALVLDLRANNVAKRSTMPKVFGEKRFNSIFEGGPELMOPH
VIDQAINSAINYLGAEISRLPTQSTPTSSDMGVMSMYPLRPPAPAGHLSADSAE
NLLLLAKSSASSEKDGSPSHGQSTDTESNNEKAGVAGSLIYLTNHTISGVNRG
VLPLVKEEQROVEAMRASIEIASSEGFVLSGEQVRAVRCHECRILFLDHVMYTIH
MCHGFRDPFECLGHRSDRYERSSHTRGEHRY"
BASE COUNT 638 a 517 c 488 g 539 t 1 others
ORIGIN

Query Match 26.4%; Score 165.8; DB 4; Length 2183;
Best Local Similarity 64.9%; Pred. No. 1.8e-37;
Matches 276; Conservative 0; Mismatches 122; Indels 27; Gaps 1;

Qy 229 ggtgaacgcccattccagtgtaacagtggtgggcatcttttactcagaaggttaacctc 288
Db 212 GGAGAGCGTCCATCCAGTGCACCCAGTGTGGCGCTCTTTACCCAGAGAGGCAACCTG 271

Qy 289 ctccgcccattaaactgcacacaggggaaaaaaccttttaagtgcaccttgcaactat 348
Db 272 CTACGTCAACATCAAGTCCCACTCAGGAGAGAGCCCTTCAAGTGTCACTGTGCAACTAT 331

Qy 349 gcatgccaagaagagatgcgtccacgggcatcttaggcacacattctgtggagaacccc 408
Db 332 GCCTGTGCGCGAGAGAGCGCCCTCAGCGGTACCTGCGGTACCCACTCTGTGTGAAAACCC 391

Qy 409 tacaatgtgagtttgggaaggagtacaagcagagaagttcccttgaggagcacaaag 468
Db 392 CACAAGTGTCTTACTGTGGCGGAGCTACAGCAGCGTAGCTCTCTGTGAGGAGACACAAG 451

Qy 469 gaggcgtgcgtacattcttcagagacactgacccaggggacactgcaagtgc----- 522
Db 452 GAGCGGTGTCACAACTACCTCCAGTGCATGGGGCTCAGAACACAGCATCTATACAGTAGTA 511

Qy 522 -----ggaggcaagacacatcaaacgacagatggaagtgaaga 561
Db 512 AAGGAAGAAAGCAACCAAGATGAGCAGAGGGAAGACTTAAAGCCAGATGGGATCTTAAGA 571

Qy 562 gctctcgtagtgagcagattagcaagcaatgtggcaaaacgaaaagctcaatgcctcag 621
Db 572 GCCTTGTGTAGACACACTAGCTAATAATGTAGCCAAACGTAAGACCACTATGCCACAG 631

Qy 622 aaatt 626
Db 632 AAGTT 636

RESULT 14
OMU92199
LOCUS OMU92199 2079 bp mRNA VRT 02-MAY-1997
DEFINITION Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, Ik-8 isoform,
complete cds.
ACCESSION U92199
NID g2062737

```



```

QY 134 gtacagggcccaaccagtggaagagtaactgcgatgtgtgtgattatccctgcatacagct 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1045 GCATCCGGCTACCCACAGCGCAAGCTCAAGTCGACGCTCGGCGATGGTCTGCTGATTGGC 1104

QY 194 tcaatgtcttaagtgttcataaagcgaagccatactggtgaagcccaattccagtgtaatc 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1105 CCAATGTGCTCATGGTACAAAGCGCAGCCACACTGGGAGAGGCGCTTCCACTGTAAATC 1164

QY 254 agtgtggggcatcttttactcagaagaagttaacctctcccccacattaaactaacacacag 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 AGTGTGTGCTCTCTTACACAGAGAGGGCAATCTGCTTCGCCACATCAAGCTGCACCTGG 1224

QY 314 gggaaaaacattttaagtgtcaacctctgcactatgcataaggaagagatgcgtca 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1225 GGGAGAAAGCCCTTCAAGTGCCCTTCTGCAACTATGCTGCGCGCGGTGAGCGCACTCA 1284

QY 374 cggggcatcttaggacacatt-----ctgtggagaaacccctacaattgtg 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1285 CTGCCACCTTCGACACACTCAGTCTCTCCGCCACCGTGGGCAACCTACAAGTGCA 1344

QY 419 agttttggaagaggttacaagcagagaagttcccttgaggagcacaagagagcgctgcc 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1345 ACTACTGTGCGCGAGGTACAAACAGCAAGTACCTTGGAGGACACAAAGAGGTGCC 1404

QY 479 gtacattcttcagag 494
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 ARAACTACCTACAGAG 1420

RESULT 11
MUSIKAROS 1550 bp mRNA ROD 11-MAY-1995
LOCUS Mouse Ikaros DNA binding protein (Ikaros) mRNA, complete cds.
DEFINITION L03547
ACCESSION g198286
NID L03547.1 GI:198286
VERSION DNA-binding transcription factor; Ikaros; Ikaros DNA binding
KEYWORDS protein; transcription; zinc finger protein; zinc-finger
transcription factor.
SOURCE Mus musculus (library: Lambda ZAP EL4) adult cDNA to mRNA.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1550)
AUTHORS Georgopoulos,K., Moore,D.D. and Drefler,B.
TITLE Ikaros, an early lymphoid-specific transcription factor and a
JOURNAL Science 258 (5083), 808-812 (1992)
MEDLINE 93068267
FEATURES
    source
        1..1550
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /cell_type="T cell"
            /dev_stage="adult"
            /tissue_lib="Lambda ZAP EL4"
            39..1334
            /db_xref="GI:198287"
            /gene="Ikaros"
            39..1334
            /gene="Ikaros"
            /codon_start=1
            /product="Ikaros DNA binding protein"
            /protein_id="AAA66193.1"
            /db_xref="PID:g198287"
            /db_xref="GI:198287"
            /translation="MDVDEQDMQSVSGKSPVSDTPDGDPEMPVPEDLSTSCAQ
            QNSKSDRGGERPFQNCQASQTQGNLLRHKLKSGEPPFKKCLNCVACRRRDALT
            GHURTSVGRPHKCYGRSYKORSSLEHKERCHNYLESMLPGVCPVKEETNNE
            HMDLCKINAEKSLVLDRLASNAVKKRSMQPKFLGDKLSDMPYDSANYEKEDMTS
            HVMDQAINAIYLGAEISRLVPTPPGSEVPVTSVYQLHKKPPSDGPPRNSHQAQ
            DAVNLLLSKAKSVSREASPSNSQDSTDTESNAEORSGLIYLTNHNPHARG
            LALKEQRAVEVLRASENSQDAFRVVSISGEQLKYKCEHCRVFLDHDHMYTHMG

```

```

BASE COUNT 385 a 431 c 423 g 311 t
ORIGIN
HCHGFRDPFECNMGYHSQDRYEFSSHITRGEHRYHLS"

Query Match 27.3%; Score 171.6; DB 12; Length 1550;
Best Local Similarity 67.3%; Pred. No. 3.8e-39;
Matches 284; Conservative 0; Mismatches 114; Indels 24; Gaps 2;

QY 229 ggtgaagcccatccagtgtaactcagtggtgggcatcttttactcagaaagtaacctc 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 GGTGACAGCGCTTTCAGTGCAACACAGTGTGGGCGCTCTTTACCCAGAAAGCGACCTC 257

QY 289 ctccgccattaaactgcacacagggggaacaccttttaagtgtcacctctctgcaactat 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 CTGCGGCACATCAAGCTGCACCTCGGTGAGAGCCCTTCAATGCCATCTTTTGCACATAT 317

QY 349 gcatgccaagaagatgcgtccacagggggcatcttaggacacattctgtgggaaaccc 408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 GCTGCGCGCGGAGGACGCCCTCACCGGCCACTCTGAGGACGCACTCCGTTGTTAAGCCT 377

QY 409 tacaaatgtgagtttgtggaagagttacaagcagagaagttcccttgaggagcacaag 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 CACAATGTGGATATTGTGGCCGAGCTATAACACAGGAAGCTCTTTAGAGGAGCAFAAA 437

QY 469 gagcgtgcctgacattcttccagacactgacc-----caggggacactgcaagtgcg 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GAGCGATGCCACAACACTACTTTGGAAAGCATGTGGGCGCTTCCGGCGCTGTGCCAGTCATTAA 497

QY 523 gaggaagacacatacaag-----cagagatgggaagtgaagaagct 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 GAAGAAACTAAACCAACGAGATGGCAGAGACCTGTGTGAAGATAGGACAGAGAGGTCC 557

QY 565 ctgctactgacagattagcaagcaatgtggcaaacgaaagctcaatgcctcagaaa 624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CTTGTCTGGACAGCGCTGGCAACCAATGTGCCCAACGTAAGAGCTCTATGCGCTCAGAAA 617

QY 625 tt 626
   ||
Db 618 TT 619

RESULT 12
XL092202 415 bp mRNA VRT 02-MAY-1997
LOCUS Xenopus laevis/gilli Ikaros homolog (Ikaros) mRNA, partial cds.
DEFINITION U92202
ACCESSION g2062743
NID U92202.1 GI:2062743
VERSION Xenopus laevis/gilli.
KEYWORDS Xenopus laevis/gilli.
SOURCE Xenopus laevis/gilli.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;
Xenopus.
REFERENCE 1 (bases 1 to 415)
AUTHORS Hansen,J.D.
TITLE Isolation and characterization of Ikaros homologues in the rainbow
trout
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 415)
AUTHORS Hansen,J.D.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
FEATURES
    source
        1..415
            /organism="Xenopus laevis/gilli"
            /db_xref="taxon:8359"
            /tissue_type="thymus"
            /clone_lib="LG15.0"
            <1..>415
            /gene="Ikaros"

```


gene
CDS

/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
1..2496
/gene="Ikaros"
81..1649
/gene="Ikaros"
/note="Ik-1 isoform; similar to mouse and human Ikaros/Lyf-1; Ikaros/Lyf-1 homolog"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AA097474.1"
/db_xref="PID:g2062742"
/db_xref="GI:2062742"

/translation="MEMEAQMSQMPGRDPPNDVSEENDEAMPIDELASNLQ
HNRRDGLACINIKVEARCDENGLAIDMMNGEEDLRLVLDASGAKVNGSH
AGGPGKPYSGAGTIRPNGLKDCIGVICIGPNVLMVHRKSHGPFQCTCGGA
SFTQGNLRLHKLHSGEPKCHLCNYACRRDALSHGLRHSVKGPKCAYCGRSY
KORSILEHKECHNYLOCMGLQNSITVTVKEESNONEREDLSOMGSKRALVLDRLA
NNVARKSTMPKQFVGEKRFNSISEGGGELMOPHVIDQAINSLYLGASLRPLI
QPSSTDGMVGSMTPLHKPPAEGHLSAKDSAAENLLLSKSSASSEKDGSPSHS
GQSDTDTSNNEKAGVAGSLIYLTNHTITSGVRNGVLPVKEEQOQYEMRASIEI
ASEGFKVLSGEQVQRAYRCEHCRILFLDHVMTTHMGCHFRDPFECNLCGRHSQDR
YEFSSHMTREHRY"

BASE COUNT 707 a 592 c 597 g 599 t 1 others
ORIGIN

Query Match 33.9%; Score 212.8; DB 4; Length 2496;
Best Local Similarity 65.6%; Pred. No. 5.8e-51;
Matches 341; Conservative 0; Mismatches 152; Indels 27; Gaps 1;

QY 134 gtgcaggccacacagtggaagatgaactgcgatgtgtgtgattatctgcagct 193
Db 430 GTATCCCGCTGCCACGCGGAAGCAAGTGTGATATCTCTGGGATGTTGCAATGGCC 489
QY 194 tcaatgtcttaagtctataagcgaagccatactggtgaacccattccagtgatc 253
Db 490 CCAATGTGCTGATGGTGCAACGGAAGCCACACTGAGAGCGGTCCATTCAGTGCACC 549
QY 254 agtgggggcatcttttactcagaagaagttaacctctccgccacattaaactgcacacag 313
Db 550 AGTGTGCGCTTCTTTCACCCAGAGGCAACCTGCTACGTCACATCAAGCTCCACTCAG 609
QY 314 gggaaaaacctttaagtgtcacctctgcacactatgcattgccaaaagaagatgcctca 373
Db 610 GAGAGAGCCCTTCAAGTGTCACTGTGCAACTATGCTGTCGCGGAGAGAGCGCCCTCA 669
QY 374 cggggcatcttagacacattctgtgagaaaccctacaataatgtgattgttgggaaga 433
Db 670 CGGTACCTCGGTACCCACTCTGTGGAAACCCCAAGTGTGCTTACTGTGGGCGGA 729
QY 434 gtacaagcagagaagttcccttgagagcacaagagcgctgccgtacattttctcaga 493
Db 730 GCTACAGACGAGTAGTCTCTCGGAGGACACAGAGCGGTGTACAACTACTCCAGT 789
QY 494 gcaatgcccaggggacacatgcgaatgcgaggggca----- 529
Db 790 GCATGGGCTCCAGACACAGCATCTATACAGTAGTAAAGGAAGAAACACAGAAATGAGC 849
QY 529 --agacacataagcagagatgggaagtgaagagctctcgactggacagattagcaa 586
Db 850 AGAGGGAAGACTTAAGCCAGATGGGATCTAAGAGAGCGCTTGGTGCTAGACAGACTAGCTA 909
QY 587 gcaatgtggcaaacgaaaaagctcaatgcctcagaatt 626
Db 910 ATAATGTAGCAACAGTAAAGACACTATGCCACAGAGTT 949

RESULT 5
GGIKTRF 1557 bp mRNA VRT 03-SEP-1997
LOCUS G.gallus mRNA for Ikaros transcription factor.
DEFINITION

ACCESSION Y11833
NID 92330594
VERSION Y11833.1 GI:2330594
KEYWORDS Ikaros; transcription factor.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
Lilippo,J. and Lassila,O.
Avian Ikaros gene is expressed early in embryogenesis
Eur. J. Immunol. 27 (8), 1853-1857 (1997)
7439462
2 (bases 1 to 1557)
Lilippo,J.P.
Direct Submission
Submitted (12-MAR-1997) J.P. Lilippo, Turku Immunology Centre and
Department of Medical Microbiology, Turku University,
Kilnamyllynkatu 13, FIN-20520 Turku, FINLAND
Location/Qualifiers
1..1557
/organism="Gallus gallus"
/strain="RPL1"
/db_xref="taxon:9031"
/genome="Ikaros"
1..1557
/gene="Ikaros"
/codon_start=1
/product="Ikaros transcription factor"
/protein_id="CAA72531.1"
/db_xref="PID:e309211"
/db_xref="PID:g2330595"
/db_xref="GI:2330595"
/db_xref="SWISS-PROT:O42410"
/translation="METDEAQMDSQVSGKSPPISDVDDADEPMPVEDLSTTTGGQ
QSVKNVRLAGNIKIETQSDENGECACEDRLMDLADSGKMGSHNGPSK
AMSGVGGIRLNGKLKDCIGICIGPNVLMVHRSHGPFQCNOCGASFTOKGNL
LRHLKSGEAPFKCHLCNYACRRDALTGLRHSVCKPKKCGTGCGRSYKQRSLEE
HKERNYLOQMSISNLSYVKEETNOSEMAEDLCKIGERSLVLDRLAENARKKS
SMPQKFGKCLSDLPDATTNKEINEIMQTHVLDQAINNAISYLGAESLRPLVOTP
PVGSEVPVISPMTQLHKLPGDQNSHTAQDSAVENLLLSKASVSSERDASPSN
SCQSDTDESNNERSGLIYLTNHTGPHARGISVKEESROFDVLRAGTDSQDAFKV
ISSNGEQRVYKCEHCRVLFDLHVMTTHMGCHFRDPFECNMGYHSDRYEFSSHI
TRGEHFRHMS"

BASE COUNT 456 a 368 c 387 g 346 t
ORIGIN

Query Match 32.7%; Score 205.2; DB 4; Length 1557;
Best Local Similarity 66.5%; Pred. No. 8.8e-49;
Matches 338; Conservative 0; Mismatches 143; Indels 27; Gaps 2;

QY 146 ccagtggaagatgaactgcgatgtgtgtgattatccctgcacatcctcaatgtcttaa 205
Db 338 CTAATGGAAGCTAAATCTGATCTGTGGGATAATTTGCATCGGTCCCAATGTCTCA 397
QY 206 tgggtcataagcgaagccatactgtggaacgccattccagtgtaatacagtggtgggcat 265
Db 398 TGGTTTCACATCGAAGCCATCTACTGGAGAAGCGCCCTTCCAGTGCATAGTGGGAGCCT 457
QY 266 cttttactcagaagatgaactcctcctccgacatataactgcacaggggaaaaacctt 325
Db 458 CCTTTACGACAGAGGGAACCTCTCTCGGCCACATCAAGTTGCTCGGTGAAAGCCCT 517
QY 326 ttaagtgtcacctcttgcacattatgcataagcgaagagatgcctcagcgggcatctta 385
Db 518 TCAATGCCACCTCTCTAACTACGCTCGCGGCGACGAGGATGCCCTCACAGGCCACCTGC 577

```

Qy 386 ggacacattctgtgagaaacccctacaataatgtgagttttgtggaaggagttacaagcaga 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 815 GGACGCACTCGTGGTAAACCTCACAATATGGATATTTGGCGGAAGCTATAAACAGC 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 446 gaagtcccttgaggagcacaagagcgctgcgtacattcttcagagcactgacc--c 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 GAACGCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCCCTTC 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 504 aggggacactgcaagtgcggaggcagaacacatacaagca----- 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 CGGGCACACTGTACCCAGTCTATTAAGAGAACTAAGCAGTGAATGCGAGAAGACC 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 544 -----gagatgggaagtgaagagctctcgtactgtgacagattgaagcaagtggcaga 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 TGTGCAAGATAGGATCAGAGAGATCTCGTCTGGAGAGACTAGCAAGTAAATGTGCGCA 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 599 aacgaaaaagctcaatgcctcagaatt 626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 AACGTAAGAGCTCTATGCTCAGAAATT 1082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
HSU40462 HSU40462 3629 bp mRNA PRI 30-APR-1996
LOCUS Human Ikaros/Lyf-1 homolog (hik-1) mRNA, complete cds.
ACCESSION U40462
NID g1289370
VERSION U40462.1 GI:1289370
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3629)
AUTHORS Niefeld, W. and Meyerhans, A.
TITLE Cloning and sequencing of hik-1, a cDNA encoding a human homologue
of mouse Ikaros/Lyf-1
JOURNAL Immunol. Lett. 49 (1-2), 139-141 (1996)
MEDLINE 96252222
REFERENCE 2 (bases 1 to 3629)
AUTHORS Niefeld, W.
TITLE Direct Submission
Submitted (10-NOV-1995) Wilfried Niefeld, Department of Virology,
University of Freiburg, Institute for Medical Microbiology and
Hygiene, Hermann-Herder-Strasse 11, Freiburg 79104, Germany
LOCATION/Qualifiers
FEATURES
source 1. .3629
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
169. .1728
/gene="hik-1"
169. .1728
/gene="hik-1"
/feature="hik-1"
/note="similar to mouse Lyf-1, encoded by GenBank
Accession Number S74708; similar to mouse Ikaros
DNA-binding protein, Swiss-Prot Accession Number Q03267"
/codon_start=1
/product="Ikaros/Lyf-1 homolog"
/protein_id="AAC50459.1"
/db_xref="PIR:g1289371"
/db_xref="GI:1289371"
/translation="MDADEGDMQSVKSPVSDPTDEGDEPMPIDELSTTSGQ
OSSKDRVASNVKVEVTSQDEENGACEMGECAEDLRMLDASGEKMGSHRDGSS
ALSGVGIRLPNGKLDICIGIPNLMVHKRSHTGERPFQCCGASFTQKGNL
LRHKLSGSEKPFKCHLCNACRRDALTGLHRTSVKGVKPKGCGYKQKSLLE
HKEKHNLSMGLPGTILYPIVKEETHSEMAEDLKIGERSLVLDRLASVAKRKS
SMPQKFLDKGLSPDYSSAYEKENEMKSHYMDQAINNAINYLGAESLPLVPTP
PGSEVPVLTSPMYQLKPLAEGTPRSHSAQSDAVENLLSKAKLVPSEKASPN
SCQDSTSNNEQSRGLIYLTNHIAPHARGLSLKEEHRAYDLLRAAENSODALR
VYSTSGEQMKYKCEHCRVFLDHWMTIHMGCGRDPECNMGYHSQDRYEFSSH
ITRGEHRFMS"
BASE COUNT 917 a 1002 c 936 g 773 t 1 others

```

```

ORIGIN
Query Match 34.2%; Score 214.8; DB 11; Length 3629;
Best Local Similarity 67.7%; Pred. No. 1.5e-51;
Matches 344; Conservative 0; Mismatches 137; Indels 27; Gaps 2;

Qy 146 ccagtgaagaagatgaactgcgagtgtgtgtgattatctctgcatacagattcaatgtcttaa 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 CTAACGGAATAACTAAAGTGTGATATCTGTGGGATCATTTTGCATCGGGCCCAATGTGTCTCA 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 206 tgggttaagcagaagcaccatactactgtggaagccatccagtgtaatacagtggtgggcat 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 TGGTTCACAAAGAAGACACACTGGAGAACGGCCCTTCCAGTGCACATCAGTGGCGGGCCT 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 266 cttttactcagaaggttaacctctccgccacattaaactgcacacagggggaacacctt 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 CATTCACCCAGAGAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGGGGAGAGGCCT 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 326 ttaagtgtcaacctctgcaactatgcatgccaagaagaagatgcgctcacgggggcatetta 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 TCAAATGCCACCTCTGCAACTACGCTGCGCGGAGGAGCGCCCTCACTGGCCACCTGA 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 386 ggacacattctgtggaagaacccctacaataatgtgagttttgtggaaggagttacaagcaga 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 GGAGCGCACTCGTGGTAAACCTCACAATATGGATATTTGGCGGAAGCTATAAACAGC 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 446 gaagtcccttgaggagcacaagagcgctgcgtacattcttcagagcactgacc--c 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 GAAGCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCGCCTC 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 504 aggggacactgcaagtgcggaggcagaacacacacatcaaaagca----- 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 CGGCGCACACTGTATCCAGTCAATTAAGAGAACTAATACAGTGAATGCGAGAAGACC 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 544 -----gagatgggaagtgaagagctctcgtactggcagattgaagcaagtgtgcca 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 TGTGCAAGATAGGATCAGAGAGATCTCTCGTCTGGAGAGACTAGCAAGTAAATGTGCGCA 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 599 aacgaaaaagctcaatgcctcagaatt 626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 AACGTAAGAGCTCTATGCTCAGAAATT 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
OMU92201 OMU92201 2496 bp mRNA VRT 22-JAN-1998
LOCUS Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, 1x-1 isoform,
complete cds.
ACCESSION U92201
NID g2062741
VERSION U92201.1 GI:2062741
KEYWORDS rainbow trout.
SOURCE Oncorhynchus mykiss
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Hansen, J.D., Strassburger, P. and Du Pasquier, L.
TITLE Conservation of a master hematopoietic switch gene during
vertebrate evolution: isolation and characterization of Ikaros from
teleost and amphibian species
JOURNAL Eur. J. Immunol. 27 (11), 3049-3058 (1997)
MEDLINE 98056818
REFERENCE 2 (bases 1 to 2496)
AUTHORS Hansen, J.D.
TITLE Direct Submission
Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
JOURNAL
FEATURES
source 1. .2496
/organism="Oncorhynchus mykiss"

```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 06:39:34 ; Search time 701.47 Seconds
(without alignments)
2847.213 Million cell updates/sec

Title: US-09-019-348-7
Perfect score: 628
Sequence: 1 gaaagagatgagaattttt.....ctcaatgcctcagaattca 628

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_vi.*
- 18: em_fun.*
- 19: em_hlg.*
- 20: em_hum1.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_vi.*
- 34: gb_hlg1.*
- 35: gb_hlg2.*
- 36: gb_in1.*
- 37: gb_in2.*
- 38: em_ba1.*
- 39: em_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description
1	496.8	79.1	1521	12	AF001293 Mus muscu

ALIGNMENTS

RESULT 1

AF001293
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM

AF001293
Mus musculus
Mus musculus
92150043
AF001293.1
house mouse
house mouse
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 1521)
Morgan, B., Sun, L., Avital, N., Andrikopoulos, K., Ikeda, T.,
Gonzales, E., Wu, P., Neben, S. and Georgopoulos, K.
Alolols, a lymphoid restricted transcription factor that interacts
with Ikaro to regulate lymphocyte differentiation
EMBO J. 17, 2004-2013 (1997)
2 (bases 1 to 1521)
Morgan, B.
Direct Submission
Submitted (24-APR-1997) CBRC, MGH East, Bldg. 149 13th Street,
Charlestown, MA 02129, USA

FEATURES

2	216.4	34.5	1788	10	S80876	S80876 IKAROS-hik1
3	214.8	34.2	3629	11	HS040462	U04062 Human Ikaro
4	212.8	33.9	2496	4	OM092201	U92201 Oncorhynch
5	205.2	32.7	1557	4	GGIKTRF	Y11833 G.gallus mR
6	204	32.5	2301	4	OMU92198	U92198 Oncorhynch
7	197	31.4	684	12	S74708	S74708 Ikaro/LyF-
8	190.4	30.3	1847	12	AF044257	AF044257 Mus muscu
9	180.2	28.7	2309	4	AF092175	AF092175 Danio rer
10	176	28.0	2688	12	AB017615	AB017615 Mus muscu
11	171.6	27.3	1550	12	MUSIKAROS	L03547 Mouse Ikaro
12	166.2	26.5	415	4	XLU92202	U92202 Xenopus lae
13	165.8	26.4	2183	4	OMU92200	U92200 Oncorhynch
14	158.6	25.3	2079	4	OMU92199	U92199 Oncorhynch
15	104.6	16.7	148640	4	AFU56116	AFU56116 Fugu rubr
16	103.2	16.4	2582	9	HUMZIFI	L32163 Homo sapien
17	103.2	16.4	4208	11	AF011573	AF011573 Homo sapi
18	103.2	16.4	2582	14	G28565	G28565 human STS S
19	93.2	14.8	1799	12	MMU224805	AJ724805 Mus muscu
20	93	14.8	586	4	AF024439	AF024439 Xenopus l
21	92.4	14.7	3369	10	HS011806	AJ011806 Homo sapi
22	92.4	14.7	3186	10	HSOZF	X70394 H.sapiens O
23	92	14.6	1434	12	MMNZFP	X52533 Mouse mRNA
24	92	14.6	3492	12	MMZFP3P1	X89264 M.musculus
25	92	14.6	2849	12	MUSZFPAA	D45210 Mouse mRNA
26	91.6	14.6	3923	10	HSU57796	U57796 Human zinc
27	91.6	14.6	4848	11	HUMLD512P3	U80800 Human zinc
28	91.4	14.6	128779	9	HS29K1	298745 Human DNA s
29	91.4	14.6	187543	11	AC005678	AC005678 Homo sapi
30	91.4	14.6	169931	11	AC005822	AC005822 Homo sapi
31	90.6	14.4	5471	9	AB007886	AB007886 Homo sapi
32	90.4	14.4	1949	12	MMZFPR	X64413 M.musculus
33	90.2	14.4	2342	12	MMZFPP35	X17617 Mouse 2fp-3
34	90.2	14.4	2268	12	MMZFPPB3	M36146 Mouse zinc
35	89.6	14.3	1414	10	HSK0X1	X52332 Homo sapien
36	86.6	13.8	1944	9	HUMZNF8	M29381 Human zinc-
37	86.6	13.8	3743	11	AF020591	AF020591 Homo sapi
38	86	13.7	2615	12	MMKR2	Y00850 Murine mRNa
39	86	13.7	819	12	MUSFPB	M15709 Mouse finge
40	85.8	13.7	159506	9	HS3418	AL021918 Homo sapi
41	85.6	13.6	3376	3	BT02FGENE	X81804 B.taurus Oz
42	85.6	13.6	44544	11	AC004262	AC004262 Homo sapi
43	85.6	13.6	2002	12	AF045565	AF045565 Mus muscu
44	85.2	13.6	166711	11	AC006116	AC006116 Homo sapi
45	84.8	13.5	3585	42	AF153201	AF153201 Homo sapi


```
Db 290 GGPBARCKCGKVKYNNHFLAIHQSHGTGPRPPKCNKCGKFAQKSLQVHTRMTGPRP 349
QY 173 FKCHLCNYA-----
Db 350 YTCVCSKALTTKHSLLSHGOKSFTCDGCKGYFSONRQLKSHYRVHTGSHLPECK 409
2, 182 -CORR-----DALTGHLRTHSVKPYKCEGCGRSYKORSSLEEHKEKRCFLQNP----- 231
Db 410 DCHRFMDYSQKHLRTHGTGKPFTEICGKSFTKSLQTHIRHSG--EKPYSCGIC 467
QY 231 --DLGDAASVEARHKAEMGSE-----RALVLDRLASNAVKRKKSSMPQKFIKGRH 279
Db 468 GKSFSD--SSAKRHCHILHTGKKFFCPCPCNQLQFARDNLKAHL-----KIHSKEKH 517
QY 480 CPDANYNPGMYRKENB---MMQTRMMDQAINNAISYLGAFAFRPLVQTP--PAPTSEMP 335
Db 518 ASDASSISGSNTEEVNRNQLQPOLSTSGEQEIQLLVDSVHNINFMPPGSGISIVT 577
QY 336 VISSVYPIALTRADMPGAPQEMKRRILLPEKILPSE--RGLSPNNSAQDSTDTSNH-- 393
Db 578 AESQNNMTADQANLTLTQOPEQLNLILSAQOQTEHTQSLNMIESQMGPSQTEPVHV 637
QY 393 -----EDROHLYQO-----SHVLPQARNGMPLLKVEPRSFELLKPPP 430
Db 638 ITLSKETLEHLHAHQEOTEELHLATSTSDPAQHLQLTQEPGPPPTHHVP-----QPTP 691
QY 431 I 431
Db 692 L 692

RESULT 11
O43724
ID O43724 PRELIMINARY; PRT; 1207 AA.
AC O43724;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE D29K1.2 (FRAGMENT)
GN D29K1.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA TUBBY B.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98745; CAB11428.1;
DR PFAM; PF00096; zf-C2H2; 19.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 1207 AA; 140376 MW; C54700DD CRC32;
```

```
Query Match 11.1%; Score 302.5; DB 4; Length 1207;
Best Local Similarity 21.0%; Pred. No. 3.6e-16;
Matches 111; Conservative 71; Mismatches 207; Indels 139; Gaps 15;

QY 4 IQPTVELKSTEEQPLTPESPDALNDYSLPKPHEIENVDSREAPANEDEDGEDSMKVKE 63
Db 744 LOETVLRKKEGSPMSLQSKAQPKYESPE-----LESQOQDVETNGYGNLKQE 794
QY 64 YSDR-----DENIMKPEPMGDAEE-----SEMPYSARE-----YGDYESIKLER- 104
Db 795 VSEMEPHGTSSKFNDMSKSARCGTREPETEPEPSACSREDKQPTCENGVSILTEN 854
QY 104 --HVPYNSRPTSGKMNCDVCGLSICISFNVLVHVKRSHGTGERPFQCCGASFTQGNLL 161
Db 855 SDHTEHORICPGEESYCDGCGKAFSQSHSLIEHQRIHTGDRPYKCECGKAFGRVLI 914
QY 162 RHILKHTGKPFKCHLNYACORDALTGHLRTHSVKPYKCEGCGRSYKORSSLEEH-- 220
```

```
Db 915 RHKIIHTGKPYKCNKCGKAFGRWSALNQHRLHTGKHYHCNDCGKAFSQKAGLFHHK 974
QY 220 ---KERCRAFTQNPDLGDAASVEARHKAEMGSE-----RALVLDRLASNAVKRKS 268
Db 975 IHTRDKPYQCTCNKFSRSLTTOHQVHTGAKPYECNCKGKAFYN-----SSL 1025
QY 269 MPQKFIGKRRHCFDANYNPGMYKENEMMOTRMMDQAINNAISYLGAFAFRPLVQTPPA 328
Db 1026 VSHQBIHHEKCYQC-----KCGKSFQSG-----LIQHORI 1058
QY 329 PTSEMPVPISSVYPIALTRADMPGAPQEMKRRILLPEKILPSEGLSPNNSAQDSTD 388
Db 1059 HTGK-PYKCDVCEKAF-----IORTSEHQRIHTGDRPYKCDKCGKRAFTQR 1105
QY 389 DSNHEDROHLYQOQSHVLPQARNGMPLLKVEPRSFELLKPPPICLRDSIKVINKEGYMD 448
Db 1106 SVLTE-----HQRIHT-----GE-R 1119
QY 449 VPRCDCHVFLDYVMFTIHMCHGCFRDPFECNMGCGYRSHDRYFSSH 496
Db 1120 PYKDECGNAFRGITSLIQHORIHTGKPYQCGKAFGRQSRDLSKH 1167

RESULT 12
O43361
ID O43361 PRELIMINARY; PRT; 691 AA.
AC O43361;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE R30217.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA LAMERDIN J.E., MCCREADY P.M., SKORONSKI E., ADAMSON A.W.,
RA BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STILLMAGEN S.,
RA GARNES J., DANGANAN L., POUNDSTONE P., CHRISTENSEN M., GEORGESCU A.,
RA AVILA J., LIU S., BRUCE R., QUAN G., MONTGOMERY M., OW D., NOLAN M.,
RA TRONG S., KOBAYASHI A., OLSEN A.O., CARRANO A.V.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004076; AAB97932.1;
DR PFAM; PF00096; zf-C2H2; 16.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 12.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 691 AA; 80606 MW; 3FEA0BAF CRC32;
```

```
Query Match 11.1%; Score 301.5; DB 4; Length 691;
Best Local Similarity 24.2%; Pred. No. 2.1e-16;
Matches 100; Conservative 60; Mismatches 152; Indels 101; Gaps 16;

QY 110 SRPTSGKMNCDVCGLSICISFNVLVHVKRSHGTGERPFQCCGASFTQGNLLRHKLHTG 169
Db 206 SRPT--PYECTOCGKAFLTQAHLVGHQKHTGTGEQPYECNCKGKFFMYNSKLRHOKRVHG 263
QY 170 EKPFGKCHLNCYACQRRDALTGHLRTHSVKPYKCEGCGRSYKORSSLEEHKEKRCRA---F 226
Db 264 ERYECSECGKLFMDSTLGRHQRVHTGERPFECSCGKFFSHRSTLNNH--QRVHAGKRL 322
QY 227 LQNPDLGDAASVE---ARHKAEMGS-----ERALVLDRLASNAVKRKKSSMPQKFTG 275
Db 323 YKCEGKAFSLKHNHVQHLKIHTGERPYECCEKAFV-----RKSHLVQH---- 370
QY 276 EXRHCFDANYNPGMYKENEMMOTRMMDQAINNAISYLGAFAFRPLVQTPPAPTSEMP 335
Db 370 QKIHT-DA-----FSKRSDLIQHKRID-----INRPYTCSECK 403
QY 336 VISSVYPIALTRADMPGAPQEMKRRILLPEKILPSEGLSPNNSAQDSTDTSNHEDR 395
Db 404 AF-----LTQA-----HLVGHQKIHTGER-----PYECTOCAKAFVRK 436
```

```
CC Clogmia albipunctata (mothmidge).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Diptera; Nematocera; Psychodoidea; Psychodidae; Clogmia.
CC [1]
RP SEQUENCE FROM N.A.
RA ROHR K.B., TAUTZ D., SANDER K.;
RT "Segmentation gene expression in the mothmidge Clogmia albipunctata
RT (Diptera, Psychodidae) and other primitive dipterans.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131041; CAA10281.1; -.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 485 AA; 55367 MW; 01F6DECA CRC32;

Query Match 12.6%; Score 342.5; DB 5; Length 485;
Best Local Similarity 23.9%; Pred. No. 6.7e-20;
Matches 127; Conservative 69; Mismatches 171; Indels 165; Gaps 23;

QY 52 DAGESMKVKDYSDRDENIMKP-----EPMGDAESEMYPYAREYSDYESIKLER 103
DB 37 DSGEDS-----HSPMPSDLLEPVITDADYDDENDAEEDD-----DIRTPKINS 80
QY 104 HVPYDNRPTSGKM-----NCDVCGLSGISFNVLVWVKRSH-TGERPFCNOCCASFTQ 156
DB 81 H-----GKMTYKCKOCDFATVTKLSF---WEHNRTHIKPEKMLKCKQCPFITEY 127
QY 157 KGNLLRHKLHTGEKPFKCHLNCACORRDALTGHLRTHSVKEPKYKCFGCRSRYKORSLL 216
DB 128 KKHLEYHLRNHNGSKPFCQKQCNVSCVKNLSHSHKSHSNLYQYRCKDCNVATKYCHSL 187
QY 217 EEHKERCRAFTLQNDL-----GDAASVEARHIKAEMGSEALVLDRLA 259
DB 188 KLRH---LRKYSHPNPVNLNYDGTNPRLRIIDVYGRRGPKVKFKHDEGGH-----NLLN 238
QY 260 SNAVKRKSMPKQFFGEXRHCDFDANYNGYKEKENEMQMOTRMDOATNNAISYLGAFAF 319
DB 239 SNINTSRSK-----SGKRDSF-----PNF-----EOSQ----- 263
QY 320 RPLVOTPPAPTSEMPVIVSSVPIALTRADMPGAPQMEKRRILLPEK-----ILPSP 374
DB 263 ----HVPTPPSSQALAMPLNLANIQQSPMPLFYLNLFNHILLAQKAALSQISPSIN 318
QY 375 G-LSPNNSAQDST-----DTSNHDHQHLYQOSHVVYLPQARNG-----MPLL 416
DB 319 GQNEENCNEETPEKEEDPKRMSALDLSNPSTPSTVSQVKH-----KRGKRAFKLELM 373
QY 417 KEVP-----RSFELLKPPICLR-----DSIKVINKGE-----VMDVF 450
DB 374 KESSDDDEGQTIRTGLGERS-ELETPKPVQLPLTSSTTPLKTTSEDDTSVEPLQNL 432
QY 451 RCDHCHVLFLDYVMTIHMCGHFRDPFCNMGYRSHDRYEFSSHARGEH 502
DB 433 ECKFCDSFKHVLVITIMHGHYGVNDVFKCNACGKCCEDRVAFFLHIARDAH 484

RESULT 9
O14898 PRELIMINARY; PRT; 506 AA.
AC O14898:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE ZINC FINGER PROTEIN (FRAGMENT).
GN H2F6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98234550.
RA SHANNON M., STUBBS L.;
```

```
RT "Analysis of homologous XRCCI-linked zinc-finger gene families in
RT human and mouse: evidence for orthologous genes.";
RL Genomics 49:112-121(1998).
DR EMBL; AF027513; AAD12728.1; -.
DR PFAM; PF01352; KRAB; 1.
DR PFAM; PF00096; zf-C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 6.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 506
SQ SEQUENCE 506 AA; 57498 MW; 7A530D9B CRC32;

Query Match 11.3%; Score 308; DB 4; Length 506;
Best Local Similarity 29.8%; Pred. No. 4.3e-17;
Matches 64; Conservative 33; Mismatches 86; Indels 32; Gaps 3;

QY 29 YSLKPKHEIENVDSREAPANEDEGDSMKVKDEYSDRDENIMKPEPMGDAESEMYPYS 88
DB 217 FSCISHHDDNIVHKRDKVHSNSDCGDTLKVSLPTORSHTGQTYQGNCEEA----- 272
QY 89 YAREYSDYESIKLERHVPYDNRPT-----SGKMN---CDVCGLS 125
DB 272 ----ENDSSSLELHKQVHLGKSPACSTHEKDTSSYGIPVQQSVRTGKKRYWCHEGKG 327
QY 126 CISFNVLVWVKRSHITGERPFCNOCCASFTQKGNLLRHKLHTGEKPFKCHLNCYACORR 185
DB 328 FSQSSNLTQTHQVHTGKPYTCEGKSFNSOSSHLYAHLPIHTGKPYRCDSCGKGFSSR 387
QY 186 DALTGHLRTHSVKEPKYKCFGCRSRYKORSLEEKK 220
DB 388 TDLNIHVRVHTGEKPYKCEVCGKGFRTORSHLOAHE 422

RESULT 10
O43167 PRELIMINARY; PRT; 697 AA.
AC O43167:
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE KIAA0441.
GN KIAA0441.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA ISHIKAWA K., NAGASE T., NAKAJIMA D., SEKI N., OHIRA M., MIYAJIMA N.,
RA TANAKA A., KOTANI H., NOMURA N., OHARA O.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB007901; BAA23713.1; -.
DR PFAM; PF00651; BTB; 1.
DR PFAM; PF00096; zf-C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 8.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 697
SQ SEQUENCE 697 AA; 78292 MW; 20075420 CRC32;
```

```
Query Match 11.2%; Score 306; DB 4; Length 697;
Best Local Similarity 22.0%; Pred. No. 9.4e-17;
Matches 119; Conservative 69; Mismatches 177; Indels 176; Gaps 18;

QY 2 EDIOPTVELKSTEQPLTSPDALNDYSLPKPHEIENVDSREAPANEDEGADG----- 57
DB 217 EESEPTCEPSREEMPVKEK-----ENYD-----PKTEGQASQSRYSKR 256
QY 57 ----SMKVKDEYSDRDENIMKPEPMGDAESEMYPYAREYSDYESIKLERHVPYDNRSP 112
DB 257 RIWRSVXLKD-----YKLVGDQEDHGSAAK---RICGRKRKP 289
QY 113 TSGKMNCDVCGLSGISFNVLVWVKRSHITGERPFCNOCCASFTQKGNLLRHKLHTGEKP 172
RN [1]
RA SHANNON M., STUBBS L.;
```



```
Db 403 DSTDTESNEDR-----IGVGIPSSGSPQPPTIVVGRHSPAYAKEDPKPQEGLLRGT 457
QY 430 PCLDRDSIKVINKGEVMDVFRCDHCHVFLDYVFTIHMGCHGFRDPFECNMCGYRSHD 489
Db 458 PGPSKEVLRVVGESGPKAFKCEHCRIFLDHFVFTIHMGCHGFRDPFECNICGYHSQD 517
QY 490 RYEFSSHIARGHR 503
Db 518 RYEFSSHIVRGEHK 531

RESULT 4
P79751 PRELIMINARY; PRT; 417 AA.
AC P79751;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DE IKAROS-LIKE.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELLNER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF056116; AAC34387.1; -.
DR PFAM; PF00096; zf-C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 417 AA; 45664 MW; 7856E12F CRC32;

Query Match 31.8%; Score 866.5; DB 13; Length 417;
Best Local Similarity 37.2%; Pred. No. 3e-62;
Matches 193; Conservative 61; Mismatches 134; Indels 131; Gaps 10;

QY 6 PTVELKSTEEQLPTSPDNLNDYSLPKPHETIENVDSREAPANEDEDAG-----EDSMKV 60
Db 7 PSIOASAEQFSGGSPSATSPNGQSSPRHSLSVHSIKVELCSDDSPGREGHPRDGGRM 66
QY 61 KDEYSDRDENIMKPEMGDAESEMPSYAREYSDYESIKLERHVPYDNRPTSCNMCD 120
Db 67 ED--GRRMEDGGRNEVGDDVDGPRGKAAGCSLSPNTASPGPI---RLPNGKLOCE 121
QY 121 VGLSCISFNVLVHKRSHGTGERPQCQCAGASFTQGNLLRHILKHTGCKPFCKLHCNY 180
Db 122 ICGIVCTGNVLVHKRSHGTGERPQCQCAGASFTQGNLLRHILKHSGEKPFKPCICNY 181
QY 181 ACORRDALTGHLRTHSVKPYKCEFCGRSYKORSLEEHKERCRAFLQNPD---LGDAAS 237
Db 182 ACRRRDALTGHLRTHAVGKAFKCYSCRSYKQOSTLEDHQRCHSYLSRQAQAAVVRQAAP 241
QY 238 VEARHIKAEMGSEALVLDRLASNAVKRKSMPQKFCIEKRCHEFDANYPGYMYEKENEM 297
Db 242 GALNMDGMNQNEKTQQVDRMTAKRKTRMPKQFLGGS----- 282
QY 298 MOTRMMDQAINNAISYLGAEAFLVQTPPTAPTSMPVIVSSVYPIALTRADMPMGAPQE 357
Db 282 -----TALGREDQPCGL--- 294
QY 358 MEKKRILLPEKILPSERGLSPNNSAQDSTDTDSNHEDR-----QHLVQQSHV 404
Db 294 -----GHAASP--NGCPDSTDETESTAEQSTRAPAQIS'NSNHLHYQ--- 335
QY 405 VLPQARNMPLLKVEPRSEFELLKPPICLRDSIKVINKEGEMVDFRCDHCHVFLDYVM 464
Db 335 -----PRG---LHSHSTARETVQVLDREGHPVRFHCHLCHILFLDHYM 376
QY 465 FTIHMGCHGFRDPFECNMCGYRSHDRYEFSSHIARGHR 503
Db 377 FTIHMGCHGFRDPFECNCGYSSQDRYEFSSHIARGHR 415
```

RESULT 5

```
O42244 PRELIMINARY; PRT; 338 AA.
ID O42244;
AC O42244;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE IKAROS-RELATED TRANSCRIPTION FACTOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA TURPEN J., KELLEY C., MEAD P., ZON L.;
RL Immunity 0:0-0(1997).
DR EMBL; AF024439; AAB81380.1; -.
DR PFAM; PF00096; zf-C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 2.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
FT NON_TER 328
FT NON_TER 328
SQ SEQUENCE 328 AA; 37116 MW; CCC9E71E CRC32;

Query Match 25.0%; Score 681.5; DB 13; Length 328;
Best Local Similarity 44.4%; Pred. No. 1.9e-47;
Matches 147; Conservative 55; Mismatches 100; Indels 29; Gaps 8;

QY 166 LHTGKPFKCHLNCACRRDALTGHLRTHSVKPYKCEFCGRSYKORSLEEHKERCRA 225
Db 1 IRLGKPFKPCFCSYACRRDALTGHLRTHSVGPKHCNCGRSYKORSLEEHKERCCHN 60
QY 226 FLQNPDLGDAASV-----EARIKAEMGS-----ERALVLDRLASNAVKRKSMP 270
Db 61 YLQNVGMEAGGIVHASPMDCEKNPMENNNMSPFEPAPVIERLASNMGRKKSSTP 120
QY 271 QKFTIGK--RHCF--DANYNPGYMEKENEMMOTRMMDQAINNAISYLGAEAFLVQTPP 327
Db 121 QRFGLGKLMRYGYPDLHFD--MAYEKEAETIQSQMMDQAINNAITYLOADALRLIHHSA 178
QY 328 APTSEMPVIVSSVYPIAL--TRADMPMGA-----POEMEKKRILLPEKILPSERGLSPN 379
Db 179 AAMPEVPIVSSLYSQVYHPARVERPTSRETSDSNNDMDGPISLRPNKHPQEREASPS 238
QY 380 NSAQDSTDTDSNHEDRQHLVQQSHVLPQARNMPLLKVEPRSEFELLKPPICLRDSIKV 439
Db 239 NSCLDTTDSSESSHE--APAILPFGNALNSKRKQSSVYPKHEGKPLDATKSSLGSSKDYVR 297
QY 440 INKEGEMVDFRCDHCHVFLDYVFTIHMG 470
Db 298 FNSEGEQIRAFKCEHCRIFLDHFVFTIHMG 328

RESULT 6
O13100 PRELIMINARY; PRT; 138 AA.
ID O13100;
AC O13100;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1) (FRAGMENT).
OS IKAROS OR LYF-1.
OS Xenopus laevis/Gillii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98056818.
```

361	Qy	KRILLPEKILPSERGLSPNNSAQDS	TDTSDNHEDRQHL	YQQSHVVL	PQARN	GMP	LLKEVP	420
361	Qy	KRILLPEKILPSERGLSPNNSAQDS	TDTSDNHEDRQHL	YQQSHVVL	PQARN	GMP	LLKEVP	420
421	Qy	RSFELKPPPICLRDSIKVINKEGEVMD	VFRCDHCHVFLD	YVMTIHMGC	HGCFRDP	PEC	480	
421	Db	RSFELKPPPICLRDSIKVINKEGEVMD	VFRCDHCHVFLD	YVMTIHMGC	HGCFRDP	PEC	480	
481	Qy	NMCGYRSHDRYEFSSHIARGE	HRAMLK	507				
481	Db	NMCGYRSHDRYEFSSHIARGE	HRAMLK	507				
RESULT		2						
O93581	ID	O93581	PRELIMINARY;	PRT;	537	AA.		
AC	AC	O93581:						
DT	DT	01-NOV-1998 (TReMBLrel. 08, Created)						
DT	DT	01-NOV-1998 (TReMBLrel. 08, Last sequence update)						
DT	DT	01-NOV-1999 (TReMBLrel. 10, Last annotation update)						
DE	DE	IKAROS.						
OC	OC	Brachydanio rerio (Zebrafish) (Zebra danio).						
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;						
OC	OC	Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;						
OC	OC	Cyprinoides; Cyprinidae; Rasbora; Rasbora; Danio.						
RP	RP	[1]						
RN	RN	SEQUENCE FROM N.A.						
RC	RC	TISSUE=SPLEEN;						
RA	RA	AMEIYA C., KAWASAKI H.;						
RT	RT	"Characterization of zebrafish ikaros, a gene necessary for						
RL	RL	differentiation of the immune system.";						
RT	RT	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.						
DR	DR	EMBL; AF092175; AAC61763.1; -;						
DR	DR	FFAM; PF00096; zf-C2H2; 5.						
DR	DR	PROSITE; PS00028; ZINC_FINGER_C2H2; 5.						
DR	DR	zinc-finger; Metal-binding; DNA-binding.						
SW	SW	SEQUENCE 537 AA; 58865 MW; 6AF8330F CRC32;						

Db	410	EERSAGVSGTAATGGJTYLTNNHAPGMRNGGLPGVKEEQORHFEALRAAGMDLSIASSEG	496
Qy	437	IKYINKEGEVMDYRDCHDCHLVFLDYVMFTIHMGCHGFRDPFCNNMGYRSHDRYEFSSH	496
Db	470	FKVLSGDGEELRAYRCHHCYFLDHWYMIHMGCHGFRDPFCNLCGYRSDRYEFSSH	529
Qy	497	IARGEHR 503	
Db	530	ITRGEHR 536	
RESULT 3			
	Q9Z222	PRELIMINARY; PRT; 533 AA.	
ID	Q9Z222		
AC	Q9Z222;		
DT	01-MAY-1999 (T-EMBLrel. 10, Created)		
DT	01-MAY-1999 (T-EMBLrel. 10, Last sequence update)		
DT	01-MAY-1999 (T-EMBLrel. 10, Last annotation update)		
DE	EOS PROTEIN.		
GN	EOS.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RC	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRATN=ICR:		
RA	HOMMA Y., KIIYOSAWA H., MORI T., OGURI A., NIKAIIDO T., KANAZAWA K.,		
RA	TOJO M., TAKEDA J., TANNO Y., YOKOYA S., KAWABATA I., IKEDA H.,		
RA	WANAKA A.;		
RT	"Pos: a novel member of the Ikaros gene family expressed		
RT	predominantly in the developing nervous system.";		
EL	FEBS Lett. 0:0-0(1999).		
DR	EMBL; AB017615; BAA36213.1; -.		
DR	PROSITE; PS00028; ZINC_FINGER_C2H2; 5.		
KW	Zinc-finger; Metal-binding; DNA-binding.		
SO	SEQUENCE 533 AA; 58167 MW; 2A0315AD CRC32:		

[illegible]

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2728	100.0	507	11	O08900	O08900 mus musculus
2	1165.5	42.7	537	13	O93581	O93581 brachydactylus
3	1090	40.0	533	11	O92282	O92282 mus musculus
4	866.5	31.8	417	13	P79751	P79751 fugu rubripes
5	681.5	25.0	328	13	O42244	O42244 xenopus laevis
6	546.5	20.0	138	13	O13100	O13100 xenopus laevis
7	345	12.6	594	11	O62514	O62514 mus musculus
8	342.5	12.6	485	5	O96785	O96785 clogmia albipunctata
9	308	11.3	506	4	O14898	O14898 homo sapiens
10	306	11.2	697	4	O43167	O43167 homo sapiens
11	302.5	11.1	1207	4	O43724	O43724 homo sapiens
12	301.5	11.1	691	4	O43361	O43361 homo sapiens
13	298.5	10.9	623	11	O62510	O62510 mus musculus
14	297.5	10.9	604	4	O43309	O43309 homo sapiens
15	297.5	10.9	751	4	O60792	O60792 homo sapiens
16	296.5	10.9	591	11	P70590	P70590 rattus norvegicus
17	296.5	10.9	346	11	O62512	O62512 mus musculus
18	295.5	10.8	650	11	O62886	O62886 rattus norvegicus
19	295	10.8	693	4	Q14588	Q14588 homo sapiens
20	294.5	10.8	524	11	O88412	O88412 mus musculus
21	294.5	10.8	819	11	O92108	O92108 mus musculus
22	292	10.7	622	4	O15917	O15917 homo sapiens
23	292	10.7	273	4	O92951	O92951 homo sapiens
24	291	10.7	578	4	O15776	O15776 homo sapiens
25	291	10.7	812	11	O35483	O35483 mus musculus
26	289.5	10.6	612	4	O75802	O75802 homo sapiens
27	289	10.6	394	4	O15361	O15361 homo sapiens
28	288	10.6	367	4	Q13398	Q13398 homo sapiens
29	288	10.6	546	11	O61491	O61491 mus musculus

ZG20_XENLA
ID ZG20_XENLA STANDARD; PRT; 675 AA.
AC P18714;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GASTRULA ZINC FINGER PROTEIN XFG20-1 (XICGF20.1).
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPOIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 9004815.
RA SCHAEFER U., RAUSCH O., BOWMEISTER T., PIELER T.;
RT "Sequence-specific recognition of a repetitive DNA element by a C2H2
RT zinc-finger protein in Xenopus.";
RL EUR. J. BIOCHEM. 226:567-576(1994).
RN [2]
RP SEQUENCE OF 85-613 FROM N.A.
RX MEDLINE; 90040698.
RA NIETZEL W., EL-BARADI T., MENTZEL H., PIELER T., KOESTER M.,
RA POETING A., KNOECHEL W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. MOL. BIOL. 208:639-659(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82643; G576444; -.
DR PIR; S06565; S06565.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
DR PFAM; PF00096; zf-C2H2; 18.
DR HSP; P25490; 12NM.
DR TRANSFAC; T02366; -.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; REPEAT.
FT ZN_FING 62 84
FT C2H2-TYPE.
FT ZN_FING 90 112
FT C2H2-TYPE.
FT ZN_FING 118 140
FT C2H2-TYPE.
FT ZN_FING 146 168
FT C2H2-TYPE.
FT ZN_FING 174 196
FT C2H2-TYPE.
FT ZN_FING 202 224
FT C2H2-TYPE.
FT ZN_FING 257 279
FT C2H2-TYPE.
FT ZN_FING 286 308
FT C2H2-TYPE.
FT ZN_FING 344 366
FT C2H2-TYPE.
FT ZN_FING 373 395
FT C2H2-TYPE.
FT ZN_FING 424 446
FT C2H2-TYPE.
FT ZN_FING 452 474
FT C2H2-TYPE.
FT ZN_FING 507 529
FT C2H2-TYPE.
FT ZN_FING 535 557
FT C2H2-TYPE.
FT ZN_FING 563 585
FT C2H2-TYPE.
FT ZN_FING 591 613
FT C2H2-TYPE.
FT ZN_FING 619 642
FT C2H2-TYPE.
SQ SEQUENCE 675 AA; 77116 MW; 3468F756 CRC32;

Query Match 11.0%; Score 299.5; DB 1; Length 675;
Best Local Similarity 21.2%; Pred. No. 1.7e-13;
Matches 119; Conservative 88; Mismatches 178; Indels 175; Gaps 24;
QY 14 EEQPLTPESPDALNDYSLPK-----PHEIENVDSREAPANEDEDAGEDSMKVQDEY 64
DB 5 EEP-----YEPFNTGNNLPPYQCTGDEALSDTKSLDYLAYLEVEITDAH 53
OV 65 SDRDENIMKPEMGDAESEMPSYAREYSDESISKLHVV-PYDNSRPTSGKMNCDVCG 123
DB 54 --EESNTDKPFTC-----TECGKTTRK-PNYES-----HIRAHKGEKPFs----CMVCD 96
DB 124 LSCISFNVLMVKRSHGTGRRPQCQCQCGASFTQKGNLLRHKLHTGKPFKCHLCNACQ 183

DB 97 KAFWKSNNLVHYSVHSGEKPFSCTECDKTFNSKRAQLEKHLRVHTGKPYSCQCGKSA 156
QY 184 RDALTGLRTHSVKPYKCFGSGRSYKORSLSLEHKKRCRAFLQNPDLGDAASVEARHI 243
DB 157 HKCVLDSSHORTHTGDKPSCTECGKFSQGNLHKHLTKH--LDQP-----HL 203
QY 244 KAEMG---SERALVLDRL-----ASNVAKRSSMPQKFIGEKR--HC 280
DB 204 CAECGKTFESKSTLLEHQIHSERPLSEFGKTFSDAHNLLKHQST----FTEQKPFPC 259
QY 281 FDANYNPGMYEKENEMM---QTRMMDQAINNAISYLGAEAFRLVQVPPTSMWPVI 337
DB 260 TEC-----GEIFSNHEHLLTHOSTHEQ-----KPFCTKCGIF 295
QY 338 SSVPIALTRADMPGAPQEMEKRRILLPEKILPS-ERGLSPNNSAODSTDTDSNHEDRQ 396
DB 296 SNEHELRTHQSTHTEG-----QKSLPSTESGGTFSEHHEHLLTHQSTHTEQK 342
QY 397 HLYQOQSHVVLPOAR-----NGMPLLEKVPFRSFELLKPPICLRDSIKVINKEGVM-- 448
DB 343 H-----LPCTEGCGTFTNEQELLAHQSTHTEQKPLP--CTECGIFSDHEHLLTHQ 392
QY 448 -----DVERCDHCHVFLFDYVNFTHHGHGCHGFRDPF 478
DB 393 STHTSPSTFVGQTTEDHQSPSKDHTGKPFSCSECGKSFYKSVLKDHLVHTGKPY 452
QY 479 ECNMG-----YRSHDR 490
DB 453 HCIEGGRSYTHQSSLSKSHQR 472
RESULT 14
ZNO8_HUMAN
ID ZNO8_HUMAN STANDARD; PRT; 543 AA.
AC P17098;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 8 (ZINC FINGER PROTEIN HF.18) (FRAGMENT).
GN ZNF8.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA.
RX MEDLINE; 90169993.
RA LANIA L., DONTI E., PANNUTI A., PASCUCCI A., PENGUE G.,
RA FELICIELLO I., LA MANTIA G., LANFRANCONE L., PELICCI P.-G.;
RT "cDNA isolation, expression analysis, and chromosomal localization of
RT two human zinc finger genes.";
RL GENOMICS 6:333-340(1990).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY PRESENT IN MANY HUMAN CELL LINES
CC OF DIFFERENT EMBRYOLOGICAL DERIVATION.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29581; G340448; -.
DR PIR; B34612; B34612.
DR MIN; 194532; -.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 7.
DR PFAM; PF00096; zf-C2H2; 7.

RT repressor protein with differential DNA-binding domains.";
RL DNA CELL BIOL. 14:971-981(1995).
CC -!- FUNCTION: MAY FUNCTION AS A REPRESSOR OR SILENCER, PROTEIN, AND
CC MOST LIKELY EXERTS ITS REPRESSING ACTIVITY UPON ZINC-DEPENDENT
CC BINDING TO DNA. MAY BE INVOLVED IN PROPER SPERMATOGENESIS BY
CC REPRESSING THE EXPRESSION OF GENES UNNECESSARY OR INCOMPATIBLE
CC WITH THE MAINTENANCE OF A HAPLOID CELL STATE.
CC
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, SPLEEN, THYMUS, AND TESTIS.
CC
CC -!- DEVELOPMENTAL STAGE: THERE IS A MARKED INCREASE AFTER POSTNATAL
CC STAGES 18-20 (SIMULTANEOUSLY TO THE APPEARANCE OF HAPLOID CELL
CC STAGES). MAXIMAL EXPRESSION IS OBSERVED AROUND 2 WEEKS
CC POSTNATALLY, WITH THE EXCEPTION OF BRAIN AND TESTIS, WHERE THE
CC EXPRESSION IS HIGHEST IN EARLIER DEVELOPMENTAL STAGES.
CC
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79828; G506502; -
CC MGDI; X79828; G506502; ZFP90.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2; 13.
CC DR PFAM; PF00096; zf-C2H2; 13.
CC DR HSSP; P08047; 1SP2.
CC
CC ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN; REPEAT;
CC TRANSCRIPTION REGULATION; REPRESSOR.
CC
CC FT DOMAIN 16 53 KRAB BOX ("A BOX").
CC FT DOMAIN 54 85 KRAB BOX ("B BOX").
CC FT DOMAIN 208 628 ZINC-FINGERS.
CC FT 2N_FING 208 230 C2H2-TYPE.
CC FT 2N_FING 250 272 C2H2-TYPE.
CC FT 2N_FING 278 300 C2H2-TYPE.
CC FT 2N_FING 306 328 C2H2-TYPE.
CC FT 2N_FING 334 356 C2H2-TYPE.
CC FT 2N_FING 362 384 C2H2-TYPE.
CC FT 2N_FING 390 412 C2H2-TYPE.
CC FT 2N_FING 446 468 C2H2-TYPE.
CC FT 2N_FING 494 516 C2H2-TYPE.
CC FT 2N_FING 522 544 C2H2-TYPE.
CC FT 2N_FING 550 572 C2H2-TYPE.
CC FT 2N_FING 578 600 C2H2-TYPE.
CC FT 2N_FING 606 628 C2H2-TYPE.
CC SEQUENCE 636 AA; 72423 MW; 789C5355 CRC32;
Query Match 11.3%; Score 309; DB 1; Length 636;
Best Local Similarity 22.9%; Pred. No. 3 5e-14;
Matches 116; Conservative 50; Mismatches 140; Indels 200; Gaps 19;
QY 49 EDEDAGDSMKVDEYSDRD-ENIMKP---EPMGDAEE-SEM-----PY---- 88
DB 153 EQNKVDESSLSLTLVQLDSSIRSDCKTFGNLHNSLVTQSNILAKKPKYCDK 212
QY 88 ---SYAREYSDYESIKLERHVPYDNR---PTSGKMN--CDVCGLSGISFNLVVKRSH 139
DB 213 CRKSFHRRSLNKHETKHKGPYNSNGTDOGAQSGRKHEDCADCKTFLWRQTLEHORI 272
QY 140 TGERPFCNOCGASFTQGNLLRIKHLTKGKPKFCHLCNVCACORRDALGHLRTHSVEK 199
UB 273 TGERPFCNVCGRKAFRRSSSLGQENHATGKPYQCSLGCAGKAFORSSSLVQHQRHIGEK 332
QY 200 PYKCEFCGRSYKQRSSLEEHK-----ERCC----- 224
DB 333 PYRCNLGCRSPRHSLSLTQHEVTHSGEKPFQCKGKAFRCSSLSLVQHERHTHTGKPFEC 392
QY 224 ----RAFLNPDLDGAASVEARHIAKEMGSRALVLDRLASNAVKRSMPOKFIGEKRH 279

DB 393 SICGRAFGOSPL-----YKHMRI-----HKRSKP----- 418
QY 280 CFDAYNPGMYEKENEMQMTRMDQAINNAISYLGAFAFRPLVQTPTTAPPTSEMPVVIS 339
DB 418 -----YQSNNFSL-----AFVP--NTP----- 433
QY 340 VYPIALTRADMPGAPQEMEKRIILLPEKILPSERGLSPNNSAQD-STDTDSNHEDRQHL 398
DB 433 -----LPQG-----EGLLTVKASYHCNDCGKDFGHITDFSEHORLHA 469
QY 399 YQOSHVLVLPQANGMPLKKEVPRSFELKPPKPPICLRDSIKVINKGEVMDVFCRCHCHVL 458
DB 470 GENSY-----GSEOTLLGQQLSLHPREK-----YQCNVCGRA 502
QY 459 FLDYVMTIHMCHGCHGRDPFECNMG 484
DB 503 FKRSFIEHHRHTGKPKYECNEG 528
RESULT 10
ID ZF35_MOUSE STANDARD; PRT; 580 AA.
AC P15620.
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 35 (ZFP-35).
GN ZFP35 OR ZFP-35.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MF1; TISSUE-TESTIS;
RX MEDLINE; 90107944.
RA CUNLIFFE V.T., KOOPMAN P., MCLAREN A., TROWSDALE J.;
RT "A mouse zinc finger gene which is transiently expressed during
RT spermatogenesis.";
RL EMBO J. 9:197-205(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91065652.
RA CUNLIFFE V., WILLIAMS S., TROWSDALE J.;
RT "Genomic analysis of a mouse zinc finger gene, zfp-35, that is up-
RT regulated during spermatogenesis.";
RL GENOMICS 8:331-339(1990).
CC -!- FUNCTION: MAY ACT TO CONTROL GENE ACTIVITY DURING THE PACHYTENE
CC STAGE OF MEIOTIC PROPHASE. MAY FUNCTION AS A TRANSCRIPTION
CC ACTIVATOR.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN ADULT TESTIS.
CC -!- SIMILARITY: WITH OTHER ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17617; G55473; -
CC EMBL; M36146; G20452; -
CC EMBL; M36145; G20452; JOINED.
CC PIR; A37107; A37107.
CC PIR; S07667; S07667.
CC MGDI; MG1:99179; ZFP35.
CC PROSITE; PS00028; ZINC_FINGER_C2H2; 17.
CC PFAM; PF00096; zf-C2H2; 18.
CC HSSP; P25490; 1ZNM.
CC TRANSFAC; T00921; -
CC ZINC-FINGER; METAL-BINDING; DNA-BINDING; DEVELOPMENTAL PROTEIN;
KW

QY 140 TGERPQOCNGASFTQGNLLRLHKLHTEGKPEKCHLCNACQRRDALTGLHRLTHSVK 199
 Db 159 TGERPQOCNGASFTQGNLLRLHKLHTEGKPEKCHLCNACQRRDALTGLHRLTHSVK 218
 QY 200 PYKCFGRSQRSLSEHKEKRCRAFLQNPDLGDAASVEARHKAEMGS---ERALVLD 256
 Db 219 PYKCFGRSQRSLSEHKEKRCRAFLQNPDLGDAASVEARHKAEMGS---ERALVLD 256
 QY 257 RIASNAV---KRKSMPOKF-IGKRHCDFDANYNGYWEKENEMOTRMMDQANNAI 311
 Db 256 REAESSALMKHKKRTHSRPFCSECSRSFTNSDLTAHMKHTEFRNVNLDS----- 310
 QY 312 SYLGAERFPLVQTPAPTSSEMPVYSSYPALTRADPMGAPQEMEKRLILLPEKILP 371
 Db 310 SYLGAERFPLVQTPAPTSSEMPVYSSYPALTRADPMGAPQEMEKRLILLPEKILP 371
 QY 372 SERGLSPNNSAQDSTDTDSNHE---DROHLYQOSHVLVPOARNGMPLLKEVPRSFELL-- 427
 Db 343 SERGLSPNNSAQDSTDTDSNHE---DROHLYQOSHVLVPOARNGMPLLKEVPRSFELL-- 427
 QY 427 KPPIICLRDSIKVINKGEVMDV-----FRCDHCVLFDYVNFTHMGCHGFRDPF 478
 Db 379 ERYQCAECHKGFIQKSLVLTHTGKPEKCHLCNACQRRDALTGLHRLTHSVK 438
 QY 479 ERYQCAECHKGFIQKSLVLTHTGKPEKCHLCNACQRRDALTGLHRLTHSVK 438
 Db 439 KSCDCGKEFTQSNLILHORIHTGE 463

RESULT 8
 HUNB_TRICA STANDARD; PRT; 524 AA.
 AC Q01791;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HUNCHBACK PROTEIN.
 GN HB.
 OS TRIBOLIUM CASTANEUM (RED FLOUR BEETLE).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; COLEOPTERA; POLYPHAGA; CUCUJIFORMIA; TENEBRIONIDAE;
 OC TRIBOLIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96125163.
 RA WOLFF C., SOMMER R., SCHRODER R., GLASER G., TAUTZ D.;
 RT "Conserved and divergent expression aspects of the Drosophila
 RT segmentation gene hunchback in the short germ band embryo of the
 RT flour beetle Tribolium".
 RL DEVELOPMENT 121:4227-4236(1995).
 RN [2]
 RP SEQUENCE OF 243-311 FROM N.A.
 RX MEDLINE; 93066327.
 RA SOMMER R.J., RETZLAFF M., GOERLICH K., SANDER K., TAUTZ D.;
 RT "Evolutionary conservation pattern of zinc-finger domains of
 RT Drosophila segmentation genes".
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:10782-10786(1992).
 CC -!- FUNCTION: HUNCHBACK IS A GAP CLASS SEGMENTATION PROTEIN.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91618; G1008447; -.
 DR EMBL; L01615; G161662; -.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
 DR PFAM; PF00096; zf-C2H2; 6.
 DR HSSP; P08151; 2GLI.

KW DEVELOPMENTAL PROTEIN; GAP; ZINC-FINGER; METAL-BINDING; DNA-BINDING;
 KW REPEAT; NUCLEAR PROTEIN.
 FT DOMAIN 202 311 ZINC-FINGERS I.
 FT DOMAIN 471 523 ZINC-FINGERS II.
 FT ZN_FING 202 224 C2H2-TYPE.
 FT ZN_FING 231 253 C2H2-TYPE.
 FT ZN_FING 259 281 C2H2-TYPE.
 FT ZN_FING 298 311 C2H2-TYPE.
 FT ZN_FING 471 493 C2H2-TYPE.
 FT ZN_FING 499 523 C2H2-TYPE.
 SQ SEQUENCE 524 AA; 59514 MW; 8E8F65BD CRC32;

Query Match 11.4%; Score 310; DB 1; Length 524;
 Best Local Similarity 23.5%; Pred. No. 2.3e-14;
 Matches 124; Conservative 66; Mismatches 192; Indels 146; Gaps 20;

QY 5 OPTVELKSTEQP-----LPTESPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSM 58
 Db 112 EPLVSPKSEKEEMETTLTPCASPNNKPD-----DNQDLRLRLMSLEKSGLESS 162
 QY 59 KVK-----DEY4DRDENIMKPEPMGDAEEMSPYSAREYSDYSEIKLERHVPYDNRPTS 114
 Db 163 KTSEHSDVLSGSDN-----DAEE-----YDEQSLR-----VPKYNHSGKI 199
 QY 115 GKMDVCGGLSCIS-----FNVLMVHKRSHTGERPQCNOCCASFTQGNLLRLHKLHTE 170
 Db 200 KTFKCKQCFVAITKLEQNNHSHVIRE---DKRLTCTCPKPFITEYKHHLHYHLRNHAGS 256
 QY 171 KPFECHLCNACQRRDALTGLHRLTHSVKPYKCFGRSQRSLSEHKEKRCRAFLQNP 230
 Db 257 KPFCNKCDDYTCVNKMSLNHSHKSHNVYRSCDRCDSYATKYCHSLKILHRR----- 309
 QY 231 DLGDAASVEARHKAEMGSERALVLDLRLASNVAKRSMPOKFICEGRHCFDANYNGYM 290
 Db 309 -----YCHTENVLDE-----EGNPCPDII 328
 QY 291 YE-----KENEMQTR-MMDQAINNAISYLGAEAFRLVQTP---PAPTSEMPVYSSVYP 342
 Db 329 IDVHGTRGRPKIKTPKAAEAKPETLPFLNQQLPFGYFFGFGFPAQLLQQL----- 384
 QY 343 IALTRADPMGAPQEMEKRLILLPEKILPSEGLSPNNS-----AQDSTDTDSNHEDRQ- 397
 Db 384 --IRERQLAVGSGQ--EESRVLDSKPGCVTGEQSKRRKGPAPKVDPTQVESEDEET 439
 QY 397 --HLYQQSHVVLVPOARNGMPLLKEVPRSFELLKPPPICLRDSIKVINKEGEVMDVFRCDH 454
 Db 440 STTVFSNVVQEEAK-----KEE-----SDSNNNNKE-----EGNSCQY 475
 QY 455 CHVLFDYVMTIHMCHGFRDPFCNMGYRSHDRYEFSSHIARGEH 502
 Db 476 CNIAFGDAVLYTIHMGYHGHFNPTCNMCGVECDKVSFFLHIAVRVSH 523

RESULT 9
 ZF90_MOUSE STANDARD; PRT; 636 AA.
 AC Q61967;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE ZINC FINGER PROTEIN 90 (ZF90-90) (ZINC FINGER PROTEIN NK10).
 GN ZF90 OR NK10.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-BRAIN;
 RX MEDLINE; 96069544.
 RA LANGE R., CHRISTOPH A., THIESEN H.-J., VOPPER G., JOHNSON K.R.,
 RA LEMAIRE L., PLOMANN M., CREMER H., BARTHELS D., HEINLEIN U.A.O.;
 RT "Developmentally regulated mouse gene nk10 encodes a zinc finger


```

FT VARSPLIC 206 246 MISSING (IN IK-4).
FT VARSPLIC 206 291 MISSING (IN IK-3).
FT VARSPLIC 247 291 MISSING (IN IK-6 AND IK-8).
SQ SEQUENCE 522 AA; 57657 MW; F39835EE CRC32;

Query Match 44.4%; Score 1211; DB 1; Length 522;
Best Local Similarity 48.5%; Pred. No. 2.5e-76;
Matches 260; Conservative 84; Mismatches 142; Indels 50; Gaps 16;

Oy 1 MEDIOPTVELKSTESOPALPTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKV 60
Db 3 MEEAQESQMGGRDPP-PNDVSEB-NDEAMPIDELS-ASSNLOHNRGDKGLACNI 58

Oy 61 KDEYSDRDINMKPEMGDAEEMSEMPYSAIEYSIKL-----ERHVPYDNS-- 111
Db 59 KVEARCDENGLAIDMMNGEEDLRL-VLDASAKVNGSHAGGPKSGPYSSAGG 117

Oy 111 -RPTSGKNCDCVGLSCISFNVLVHKSHSGTGERPFCNQCASFTQGNLLRHKLTG 169
Db 118 IRLPNKGLKDCIGVICGPNVLMVHKSHSGTGERPFCNQCASFTQGNLLRHKLTG 177

Oy 170 EKPEKCHLNCACQRRDALTGHLRTHSVKPKCFGRSKYKQSSLEHKKRCRAFLON 229
Db 178 CKPFKCHLNCACRRDALSGLRTHSVKPKCAKCAKGRSKYKQSSLEHKKRCHNYLQC 237

Oy 230 PDLGDAASV-----EARIKAEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRRH- 280
Db 238 MGLQNSIYTVVKEESNQEREDLSQMSKRALVLDRLANNVAKRKSTMPQKFIGEKRRFS 297

Oy 280 --CFDANYNPGMYKEKEMQMOTRMMDQAINNAISYLGAERPLVPTPAPTSEMVPVI 337
Db 298 NISFEG--GPG-----ELMOPHVIDQAINNAISYLGAESLRPLIQT--SPTSDMGVM 346

Oy 338 SSVYPALTRADMPMGAPOEMEKRIILL--PEKILPSEKGLSPNNSAODSTDSDNHEDR 395
Db 347 GSNYPLHPAEGHGLSAKDSNAENLLLLAKSKASSKESKSGSODSTDSDNHEEK 406

Oy 396 OH-----LYQOQSHVVLPOARG-MPLK-EVPRSFELLKPPICLRDSIKVKEGEVM 447
Db 407 AGVGASGLIYLTNHT-SGVRNGVLVKEEQORQYEAARASIAETASGFKVLSGEGQV 465

Oy 448 DVFRCDHCHVFLDYVMTIHWGCHGFRDPPECNMGYRSHDRYEFSSHARGHR 503
Db 466 RAYRCEHCRILFDHVMYTIHWGCHGFRDPPECNLCGRHSQDRYEFSSHMTGHR 521

RESULT 6
ZF37_MOUSE STANDARD: PRT: 553 AA.
AC P17141;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN 37 (ZF37) (MALE GERM CELL SPECIFIC ZINC FINGER
DE PROTEIN).
GN ZF37 OR ZFP-37.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE; 92310982.
RA BURKE P.S., WOLGEMUTH D.J.;
RT "Zfp-37, a new murine zinc finger encoding gene, is expressed in a
RT developmentally regulated pattern in the male germ line.";
RL NUCLEIC ACIDS RES. 20:2827-2834(1992).
RN [2]
SEQUENCE OF 177-553 FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE; 90301500.
RA NELKI D., DUDLEY K., CUNNINGHAM P., AKHAVAN M.;

```

```

RT "Cloning and sequencing of a zinc finger cDNA expressed in mouse
RT testis.";
RL NUCLEIC ACIDS RES. 18:3655-3655(1990).
CC -|- FUNCTION: MAY HAVE A ROLE IN REGULATING SPERMIOGENESIS.
CC -|- SIMILARITY: CONTAINS 1 KRAB BOX.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64413; G55475; -
CC EMBL; X52533; G53457; ALT_INIT.
CC PIR; S10245; S10245.
CC PIR; S22954; S22954.
CC MGI; MGI:99181; ZFP37.
CC PROSITE; PS00028; ZINC_FINGER_C2H2; 11.
CC PFAM; PF00096; zf-C2H2; 12.
CC HSSP; P08047; 1SP2.
CC ZINC_FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN; REPEAT;
CC DEVELOPMENTAL PROTEIN; SPERMATOGENESIS.
CC DONALDIN 214 544 ZINC-FINGERS.
CC ZN_FING 214 236 C2H2-TYPE.
CC ZN_FING 242 264 C2H2-TYPE.
CC ZN_FING 270 283 C2H2-TYPE.
CC ZN_FING 298 320 C2H2-TYPE.
CC ZN_FING 326 348 C2H2-TYPE.
CC ZN_FING 354 376 C2H2-TYPE.
CC ZN_FING 382 404 C2H2-TYPE.
CC ZN_FING 410 432 C2H2-TYPE.
CC ZN_FING 436 460 C2H2-TYPE.
CC ZN_FING 466 488 C2H2-TYPE.
CC ZN_FING 494 516 C2H2-TYPE.
CC ZN_FING 522 544 C2H2-TYPE.
CC CONFLICT 259 259 V -> L (IN REF. 2).
CC CONFLICT 487 487 P -> T (IN REF. 2).
CC CONFLICT 492 492 N -> K (IN REF. 2).
CC CONFLICT 531 531 F -> V (IN REF. 2).
CC SEQUENCE 553 AA; 62630 MW; BB3DF19C CRC32;

Query Match 12.6%; Score 344; DB 1; Length 553;
Best Local Similarity 23.3%; Pred. No. 1.1e-16;
Matches 114; Conservative 55; Mismatches 129; Indels 192; Gaps 16;

Oy 21 ESPDALNDYSLPKPHEIENVDSREAPA-----NEDEDAGDSMKVDEYSDRDENI 71
Db 121 KKPTANEHRKSLSHSASDVNKDEIPTRKCKDKLPNKLSDKGD-----KQTSKKCEKV 175

Oy 72 MKPEPMGDAEE---SEMPYSAIEYSIKLERHVPYDNSRPTSGK-MNCDVCGLSG 126
Db 176 CRHSASTKEDKIQTGTGRKSHCHRTPSKPE-----KAPSGKPYECNHCCKVL 223

Oy 127 ISFNVLVHKSHSGTGERPFCNQCASFTQGNLLRHKLTGKPKFCHLNCACQRRD 186
Db 224 SHKQGLLDHQRTHGTGKPYECNEGIAFSQKSHLVVQRTHGTGKPYECQCGRAHGHK 283

Oy 187 ALTGHLRTHSVKPKCFGRSKYKQSSLEH-----KRCRAFLNPDIGD 234
Db 284 ALTDHLRIHTGKPKYCKNECGKTRHSSNLMQHLRSHGTGKPYECKGKSFYNSSL-- 342

Oy 235 AASVEARHKAEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRRCHDANYNPGMYEKE 294
Db 342 -----TEHVRTHGT-----EIPY-----ECNEC----- 360

Oy 295 NEMMOTRMMDQAINNAISYLGAERPLVPTPAPTSEMVPVISSVYPALTRADMPMGA 354
Db 360 -----GKAFK-----YGSILTK----- 372

Oy 355 POEMEKRIILLPEKILPSEKGLSPNNSAODSTDSDNHEDRQHLVQOQSHVVLPOARGMP 414

```


RA HAHM K., COBB B.S., MCCARTY A.S., BROWN K.E., KLUG C.A., LEE R.,
RA AKASHI K., WEISSMAN I.L., FISHER A.G., SMALE S.T.;
RT "Helios, a T cell-restricted Ikaros family member that quantitatively
RL associates with Ikaros at centromeric heterochromatin.";
GENES DEV. 12:782-796(1998).
CC -1- FUNCTION: ASSOCIATES WITH IKAROS AT CENTROMERIC HETEROCHROMATIN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE T-CELL LINEAGE. ABUNDANT IN
CC THYMUS, LOW EXPRESSION IN BONE MARROW AND BRAIN AND
CC NO DETECTABLE EXPRESSION IN SPLEEN, LIVER, KIDNEY OR MUSCLE.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN; HELIOS A AND
CC B (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF044257; G2829277;
CC PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
CC HSP; P08047; 1SP2.
CC TRANSCRIPTION REGULATION: ACTIVATOR; ZINC-FINGER; METAL-BINDING;
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
CC DOMAIN 112 219 ZINC-FINGERS I.
FT DOMAIN 471 523 ZINC-FINGERS II.
FT ZN_FING 112 134 C2H2-TYPE.
FT ZN_FING 140 162 C2H2-TYPE.
FT ZN_FING 168 190 C2H2-TYPE.
FT ZN_FING 196 219 C2H2-TYPE.
FT ZN_FING 471 493 C2H2-TYPE.
FT ZN_FING 499 523 C2H2-TYPE.
FT VARSPLIC 111 136 MISSING (IN HELIOS A).
FT SEQUENCE 526 AA; 59388 MW; AD142BD4 CRC32;
CC
CC Query Match 45.2%; Score 1234; DB 1; Length 526;
CC Best Local Similarity 49.7%; Pred. No. 6.7e-78;
CC Matches 248; Conservative 67; Mismatches 138; Indels 46; Gaps 8;
QY 34 PHEINVDSEAPANEDEAGEDSMKVKDEYSDRDENIMKPEPMGDAESEMYSYAREY 93
DB 41 PSHMTSTNSVKLEMQSDECDROPLSREDEIRGHDEGSLEEPL--IESSEVADN--RKV 96
QY 94 SDYESIKLERHVPYDNRPTSGKMCNDCVGLCSIFSNVLMVHKRSHGTGERPFCNCGAS 153
DB 97 QDLOG-----EGGRLPNGKLKCDVCGMVCIGPNVLMVHKRSHGTGERPFCNCGAS 148
QY 154 FTQGNLLRHKLTGTEKPKFCHLCNYACORRDALTGHLRTHSVKPKCFGCGRSYKOR 213
DB 149 FTQGNLLRHKLTGTEKPKFCHLCNYACORRDALTGHLRTHSVKPKCFGCGRSYKOR 208
QY 214 SSLSEHKERCAFQNPDLGDAASVEARHTK-----AEMGSEALVLD 257
DB 209 SSLSEHKERCHNYLQNVSMGAAGVMSHHVPPMEDCKEQEPIIMDNNTLSLVPFPAVIEK 268
QY 258 LASNVAKRKSMQKFTGERKRCF---DANYNFCYMEKENMMOTRMQDAINNAISYL 314
DB 269 LTANMGRKKSSTPOKFGVKLMRFSPDIHFDMNLTKEALQSHMQDAINNAISYL 328
QY 315 GAERFLVQTPPTAPTEMVPISSVY-----PIALTRADMPGAPOMEKKRI 363
DB 329 GAELHPLMHQAPSTIAEVAPVISSAYSQVYHPNRIERPISRETSD--SHENNMDGPTS 385
QY 364 LLPKILPISERGLSPNNSAQSDTSDNHDHQHLYQOSSHVLPQARGMPLKKEVPRSF 423
DB 386 LIRPKSRPQERASPSNCSLDSDSDSSHDROS-YQGNPALNPKRKQSPAYMKEDVKAL 444
QY 424 ELKPPPICLRDSIKVINKGEVMDVRCDCHECHVFLFDYVMTFTHMGSCHGFRDPFCNMC 483

DB 445 DATKAPKSLKDIYKVFNGEGEQIRAFKCEHCRVFLDHWMTIHMCGHYRDPLECNIC 504
QY 484 GYRSHDRYEFSSHIARGHEH 502
DB 505 GYRSQDRYEFSSHIVRGHEH 523
RESULT 5
IKAR_ONCMY
ID IKAR_ONCMY STANDARD; PRY: 522 AA.
AC OL3089; OL3088; OL3086; OL3087;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIARDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;
OC SALMONIDAE; ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SHASTA; TISSUE-THYMOCYTES;
RX MEDLINE: 98056818.
RA HANSEN J.D., STRASSBURGER P., DU PASQUIER L.;
RT "Conservation of a master hematopoietic switch gene during vertebrate
RT evolution: isolation and characterization of Ikaros from teleost and
RT amphibian species.";
RL EUR. J. IMMUNOL. 27:3049-3058(1997).
CC -1- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: EIGHT FORMS OF THE PROTEIN; IK1-IK8; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT
CC OF IK1.
CC -1- TISSUE SPECIFICITY: EXPRESSION MAINLY LIMITED TO THYMUS, SPLEEN,
CC AND PRONEPHROS. VERY LOW EXPRESSION IN LIVER. NO EXPRESSION IN
CC TESTIS, BRAIN, EYE AND MUSCLE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT DAY 3-4 IN THE YOLK SAC
CC AND AT DAY 5-6 IN THE EMBRYO PROPER.
CC -1- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U92201; G2062742;
CC EMBL; U92200; G2062740;
CC EMBL; U92198; G2062736;
CC EMBL; U92199; G2062738;
CC PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
CC PFAM; PF00096; zf-C2H2; 5.
CC TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
CC DOMAIN 125 203 ZINC-FINGERS I.
FT DOMAIN 468 520 ZINC-FINGERS II.
FT ZN_FING 125 147 C2H2-TYPE.
FT ZN_FING 153 175 C2H2-TYPE.
FT ZN_FING 181 203 C2H2-TYPE.
FT ZN_FING 209 232 C2H2-TYPE.
FT ZN_FING 468 490 C2H2-TYPE.
FT ZN_FING 496 520 C2H2-TYPE.
FT VARSPLIC 55 148 MISSING (IN IK-2, IK-4 AND IK-8).
FT VARSPLIC 55 291 MISSING (IN IK-6).
FT VARSPLIC 149 291 MISSING (IN IK-5).

Db 344 ISPMQL-----HKLGNQTRSNHTAODSAVENLLLSKAKSVSSERDASFSNSQDST 398
Qy 387 DTDSDHEDRQHYQOQSHVVLPOARNGMPLLEKVPFELLKPPPTCLRDSIKVINKEGV 446
Db 399 DTSNEERSGLIYLTNIGHPHARGIS-VKEESRQFDVLRAGTDNSQDAFKVSSNCEQ 457
Qy 447 MDVFCDCRCHVFLDYVFTTHMGCHGFRDPFECNMGYRSHDRYEFSSHARGHR 503
Db 458 VRVYKCEHCRVFLDHVNYTHMGCHGFRDPFECNMGYHSDQRYEFSSHITRGEHR 514
RESULT 3
IKAR_MOUSE
ID IKAR_MOUSE STANDARD; PRT; 517 AA.
AC Q03267; Q64044; Q64045; Q64051;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS (LYMPHOID TRANSCRIPTION FACTOR LYF-1).
GN IKAROS OR LYF1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 5).
RC TISSUE-EMBRYO;
RX MEDLINE; 93068267.
RA GEORGOPOULOS K., MOORE D.D., DERFLER B.;
RT "Ikarnos, an early lymphoid-specific transcription factor and a
RT putative mediator for T cell commitment.";
RL SCIENCE 258:808-812(1992).
RN [2]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE; 95021239.
RA HAHN K., ERNST P., LO K., KIM G.S., TURCK C., SMALE S.T.;
RT "The lymphoid transcription factor Lyf-1 is encoded by specific,
RT alternatively spliced mRNAs derived from the ikaros gene.";
RL MOL. CELL. BIOL. 14:7111-7123(1994).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES EXPRESSED
CC DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: SIX FORMS OF THE PROTEIN, ISOFORMS I-VI, ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT
CC OF ISOFORM VI.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN T-CELLS AND THEIR
CC PROGENITORS, AND ALSO IN B-CELLS.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03547; G198287; -
DR EMBL; S74517; G807149; ALT_SEQ.
DR EMBL; S74518; G807151; -
DR EMBL; S74708; G807153; -
DR MGD; MGI:96535; IKAROS.
UK PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
UR PFAM; PF00096; zf-C2H2; 4.
UR HSP; P08047; 1SP2.
UR TRANSFAC; T01470; -
RW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 117 223 ZINC-FINGERS I.

FT DOMAIN 457 479 ZINC-FINGERS II.
FT ZN_FING 117 139 C2H2-TYPE.
FT ZN_FING 144 166 C2H2-TYPE.
FT ZN_FING 172 194 C2H2-TYPE.
FT ZN_FING 200 223 C2H2-TYPE.
FT ZN_FING 457 479 C2H2-TYPE.
FT VARSPLIC 53 53 M -> VAYGADGRDFPHALISDRGM (IN ISOFORMS
FT VARSPLIC 54 140 MISSING (IN ISOFORM V).
FT VARSPLIC 54 282 MISSING (IN ISOFORMS I AND II).
FT VARSPLIC 141 282 MISSING (IN ISOFORMS III AND IV).
FT CONFLICT 234 235 VC -> MY (IN REF. 2).
FT CONFLICT 480 482 MISSING (IN REF. 2).
SQ SEQUENCE 517 AA; 57336 MW; D18D2D97 CRC32;
Query Match 46.9%; Score 1279.5; DB 1; Length 517;
Best Local Similarity 52.7%; Pred No 4.8e-81;
Matches 274; Conservative 71; Mismatches 128; Indels 47; Gaps 17;
Qy 12 STEQPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGEDSMKVDEYSDRDENI 71
Db 13 SGKESPPVSDTPDE-GDEPMPVPEDLTSTGQAQNSKSDRGMSN---VKVETQSDENG 68
Qy 72 MKPEPMG-----DAESEMPTYAREYSDYESIKLERHVPYDNSRPTSGKMCDV 121
Db 69 RACEMNGEECAEDLRMLDASGERKMGSHRQGGSAALS-----GVGGIRLPNGKLKCDI 121
Qy 122 CGLSCISFNVLVHVKRSHTGERPFOCNOCGASTFQKGNLLRHKLHTGERPKFCHLCNYA 181
Db 122 CGIVCIGPNVLVHVKRSHT-ERPFOCNOCGASTFQKGNLLRHKLHSGERPKFCHLCNYA 180
Qy 182 CORRDAITGHLRTHSVKPKYKCEFCGRSYKQRSLSLEHKEKRCRAFLQNPDLGDAASY--- 239
Db 181 CRRRDALTGHLRTHSVGPKHGKGYGGRSYKQRSLSLEHKEKRCRHNLYESMGLPGVCPVIKE 240
Qy 239 EARHIK-----AEMGSRALVLDRLASNAVKRSMPOKFIGERKRCF-DANYNGPYME 292
Db 241 ETNENEMAEDLCKTAGERSLVLDRLASNAVKRSMPOKFLGDK--CLSDMPYDSA-NYE 297
Qy 293 KENEMTRMDQAINNAISVLGAEARPLVOTPPAPTSEMPVVISVYPITALTRAMP- 352
Db 298 KE-DWMTSHVNDQAINNAIYVLAESRLVQTPPG-SSEVPVVISVYHUKPPSDGPP 355
Qy 352 ---MGAPQEMEKRIPELLKIPSEGLSPNNSAQDSTDSDSNHEDRQH--LYQQSHVYL 406
Db 356 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSCODSTDSTESNAEQSGLIYLTNH-IN 414
Qy 407 POARNGMPLLEKVPFELLKPPPTCLRDSIKVINKEGVMDVPRCDHCHVFLDYVMT 466
Db 415 PHARNGL-ALKEEQRAYEVLRAASENSQDAFRVVVSTSGEQLKVYKCEHCRVFLDHVMT 473
Qy 467 IHM---GCHGRDPFECNMGYRSHDRYEFSSHARGHR 503
Db 474 IHMGCHGCHGRDPFECNMGYHSDQRYEFSSHITRGEHR 513
RESULT 4
HELI_MOUSE
ID HELI_MOUSE STANDARD; PRT; 526 AA.
AC P81183;
DT 13-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN HELIOS.
GN HELIOS.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 246-267 AND 289-306.
RC STRAIN-BALB/C; TISSUE-THYMUS;
RX MEDLINE; 98180981.

```

FT VARSPLIC 10 53 MISSING (IN IK4).
FT VARSPLIC 54 140 MISSING (IN IK2).
FT VARSPLIC 54 283 MISSING (IN IK6).
FT VARSPLIC 197 283 MISSING (IN IK3 AND IK4).
FT VARSPLIC 141 283 MISSING (IN IK5).
FT CONFLICT 11 12 QV -> FS (IN REF. 2).
FT CONFLICT 214 214 S -> T (IN REF. 2).
FT CONFLICT 245 245 N -> K (IN REF. 2).
FT CONFLICT 296 296 MISSING (IN REF. 2).
FT CONFLICT 298 298 S -> T (IN REF. 2).
FT CONFLICT 352 355 KPLA -> RRS (IN REF. 2).
FT CONFLICT 372 372 N -> Y (IN REF. 2).
FT CONFLICT 420 426 PHARGL -> RRAQV (IN REF. 2).
SQ SEQUENCE 519 AA; 57528 MW; 5A97272D CRC32;

Query Match 49.1%; Score 1339.5; DB 1; Length 519;
Best Local Similarity 53.6%; Pred. No. 3.6e-85;
Matches 279; Conservative 71; Mismatches 124; Indels 47; Gaps 14;

QY 12 STEEQPLTESPDALNDYSLRPHETIENVDSREAPANEDEDAGESMKYKDEYSDRDENI 71
Db 13 SGKESPPVSDTDE-GDEPMPIDEDLTSSGQSQSKSDRVVASN---VKVETQSDENG 68

QY 72 MKPEPMG-----DAESEMPPSYAREYSDYESIKLERHVPYDNRPTSGKMNCDV 121
Db 69 RACEMNGECAEDRLMDLDSGKMGSHRDQSSALS-----GVGGIRLPNGKLUKCDI 121

QY 122 CGLUSCSFNVLVHKRSHTGERFQCNQCGASFTQKGNLLRHKLHGTGKPFKCHLCNTYA 181
Db 122 CGIICIGPNVLMVHKRSHTGERFQCNQCGASFTQKGNLLRHKLHGTGKPFKCHLCNTYA 181

QY 182 CQRDALTGHLRTHSVKPKYKCFEGRSYKQSRSLBEHKRCRAFLQNPDL-GDAASV-- 239
Db 182 CRRDALTGHLRTHSVKPKYKCFEGRSYKQSRSLBEHKRCRAFLQNPDL-GDAASV-- 239

QY 239 -EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPKQFGEKRCRFDANNPYMYE 292
Db 242 EETNHSEMAEDLCKIGERSLVLDRLASNVAKRKSSMPKQFGEKRCRFDANNPYMYE 292

QY 293 KENEMQTRMDQAINNALSILGAERPLVQTPPTSEMPVSVIPIALTRADMPN 352
Db 301 KENEMKSHVMDQAINNALSILGAERPLVQTPPTG-GSEVPVISPMTQLHKPLAE--- 357

QY 353 GAPOE-----MEKRIILLPEKILPSERGLSPNNQAQSDTSDTNSHEDRQH--LYQOS 402
Db 357 GTPRSHSAQDSAVENLLLSLAKLVPSEREASPSNQCQSDTSDTESNNEEQRSGLIYLTN 416

QY 403 HVLVPOARNGMPLKVBPRSFELLKPPICLRDSIKVINKEGVMDVFCDCRCHVFLDY 462
Db 417 HIA-PHARGLS-LKEHRAVDLLRASENSQDALRVSTSGQMKVYKCEHRCRVFLDH 474

QY 463 VMTIHMCHGFRDPPECNMGYSHDRYEFSSHARGHR 503
Db 475 VMTIHMCHGFRDPPECNMGYSHDRYEFSSHARGHR 515

```

RESULT 2

```

IKAR_CHICK
ID IKAR_CHICK STANDARD; PRT; 518 AA.
AC 042410;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN IKAROS.
GN IKAROS OR IK.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H.B2; TISSUE=THYMUS;
RX MEDLINE; 97439462.

```

```

RA LIPPO J., LASSILA O.;
RT "Avian Ikaro gene is expressed early in embryogenesis.";
RL EUR. J. IMMUNOL. 27:1853-1857(1997).
CC -!- FUNCTION: BINDS AND ACTIVATES THE ENHANCER (DELTA-A ELEMENT) OF
CC THE CD3-DELTA GENE. FUNCTIONS IN THE SPECIFICATION AND THE
CC MATURATION OF THE T LYMPHOCYTE. ALSO INTERACTS WITH A CRITICAL
CC CONTROL ELEMENT IN THE TDT (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE)
CC PROMOTER AS WELL AS WITH THE PROMOTERS FOR OTHER GENES
CC EXPRESSED DURING EARLY STAGES OF B AND T CELL DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC HEMATOPOIETIC ORGANS
CC SUCH AS THE BURSA OF FABRICIUS, THYMUS AND SPLEEN. IN THE ADULT,
CC EXPRESSED IN SPLEEN, THYMUS, BURSA AND PERIPHERAL BLOOD
CC LEUKOCYTES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYO FROM EMBRYONIC DAY
CC 2 ONWARDS.
CC -!- SIMILARITY: BELONGS TO THE IKAROS FAMILY OF ZINC-FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y11833; E309211;
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
DR PFAM: PF00096; zf-C2H2; 5.
KW TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;
KW DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
FT DOMAIN 117 224 ZINC-FINGERS I.
FT DOMAIN 461 513 ZINC-FINGERS II.
FT ZN_FING 117 139 C2H2-TYPE.
FT ZN_FING 145 167 C2H2-TYPE.
FT ZN_FING 173 195 C2H2-TYPE.
FT ZN_FING 201 224 C2H2-TYPE.
FT ZN_FING 461 483 C2H2-TYPE.
FT ZN_FING 489 513 C2H2-TYPE.
SQ SEQUENCE 518 AA; 57586 MW; AB4DF123 CRC32;

```

```

Query Match 48.4%; Score 1321.5; DB 1; Length 518;
Best Local Similarity 52.0%; Pred. No. 6.3e-84;
Matches 279; Conservative 66; Mismatches 131; Indels 61; Gaps 13;

QY 2 EDIQTVELKSTEEQPLTESPDALNDYSLRPHETIENVDSREAPANEDEDAGESMKV 60
Db 4 DEAQDMSQVSGKSPISDVDDA--DEPMPVPEDLTSTTGGQSQVKNERVLG---NI 57

QY 61 KDEYSDRDENIMKPEPMG-----DAESEMPPSY-----AREYSDYESIKLERHVP 106
Db 58 KIETQSDENGACEMNGEACEADRLMDLDSGKMGSHUNGPGSKAMSGVGIRL----- 113

QY 107 YDNRPTSGKMNCDVCGLSCTSFNVLVHKRSHTGERFQCNQCGASFTQKGNLLRHKL 166
Db 113 -----PNGKLKCDICIGIIGPNVLMVHNRSHTERFQCNQCGASFTQKGNLLRHKL 166

QY 167 HTGKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCFEGRSYKQSRSLBEHKRCRAF 226
Db 167 HSGKPFKCHLCNYACRRRDALTGHLRTHSVGKPKYKCFEGRSYKQSRSLBEHKRCRCHNY 226

QY 227 LQNPDLG-----DAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPKQFGEK 277
Db 227 LQTMISNNLSYVKEETNQSMEDLCKIGERSLVLDRLASNVAKRKSSMPKQFGEK 286

QY 278 RHCF-DANNYPGYMEKENEMQTRMDQAINNALSILGAERPLVQTPPTSEMPV 336
Db 287 --CLSDLPYDATNTYKENEINQTHVIDQAINNALSILGAERPLVQTPPV-GSEVPV 343

QY 337 ISSVYPIALTRADMPGAPQ-----EMEKRILLPEKILPSERGLSPNNQAQSDT 386

```


THIS PAGE BLANK (USPTO)

Qy 448 -----DVFRCDHCHVLFLOYVMFTIHMGGHGERDPF 478
| | | | : | : | :
Dd 393 STHTSPTTEFGVTEDNHQSPSKDHTGKEKFFSCSECKSFFVKSVDKHLVVTGKPY 452

```

QY      479 ECNMG-----YRSHDR 490
          |  | |
Db      453 HCIEGRSYTHQSSLKSHQR 472

```

RESULT 15
S28823
finger protein ZNF43 - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1998
C:Accession: S28823
J. Lovering, R.: Trowsdale, J.
Nucleic Acids Res. 19, 2921-2928, 1991
A:Title: A gene encoding 22 highly related zinc fingers is expressed in lymphoma
A:Reference number: S26823; MUID:9127944
A:Accession: S26823
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-803 <LOV>
A:Cross-references: EMBL:X59244; NID:g38031; PID:g38032
C:Genetics:
A:Gene: GDB:ZNF43; HTF6
A:Cross-references: GDB:128653
A:Map position: 19p13.1-19p12
C:Keywords: DNA binding

Query Match	10.8%	Score 295.5;	DB 2;	Length 803;
Best Local Similarity	37.3%;	Pred. NO. 1.1e-13;		
Matches 59;	Conservative 29;	Mismatches 51;	Indels 19;	Gaps 3;

```

Qy      81  EESEMPYAREYSDYESIKLERHVPYD-----NSRPTSGKM-----NCDVC 122
      ||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      479  EECGKAFRRSSNLTKHKKHIEKK-PYKCECGKAFKWSSKLTETHTGKPYKCEEC 537

```

Oy 123 GLSCISFNVLMVHKRSHTGERPFQNCQGASFTQGNLLRIHLKHTGEKPFKCHLCNYAC 182
- : - :: ||| ||||:: : : ||| || | : |||| : || -
Dd 538 GKAFNHFSILTKHRRIRHTGSKPYKEECGGKAPTQSSNLTTTHKIHTGSKFYKCECGKAF 597

Qy. 183 QRRDALTGHLRTHSVEKPYKCEFCGRSYQORSSLEEHH 220
 : ||| : | : ||||| ||:: | : | : ||
 Db 598 TQSSNLTHKKIHTGKGKPYKCECGKAFNOFSTLTKHK 635

Search completed: November 6, 1999, 08:08:51
Job time: 71 sec

```
.
Query Match      10.9%   Score 296.5;    DB 2;    Length 510;
Best Local Similarity 31.68;  Pred.No. 5.3e-14;
Matches 73; Conservative 30; Mismatches 75; Indels 53; Gaps 7;
```

QY 6 PTVELKSTEEQLP--TESPDALNDYSLPKPHEIENVDSREAPANEDEDAGEDSM----- 59
| | : | L | L | : | L | : | L | : | L | : |
Db 309 PQVIHK---EELSLSDSGDVVN--SYV-VYAIQANCPVPAPASSGVILGTQTVPADLAH 362

QY 59 KVKDEYSRDRDENIMKKPEPM-----GDAESEMPYSYAREYSIDYESIKLHRHPYD 108
| : | : | : | : | : | : | : | : | : | : | :

A;Gene: fin
C;Keywords: DNA binding; zinc finger

Query Match 11.6%; Score 316; DB 2; Length 1350;
Best Local Similarity 24.0%; Pred. No. 7.1e-15;
Matches 121; Conservative 63; Mismatches 201; Indels 120; Gaps 20;

[illegible]

RESULT 8
I48689
gene NK10 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996
C:Accession: I48689; S49078
C:Lange, R.; Christoph, A.; Thiesen, H.J.; Vopper, G.; Johnson, K.R.; Lemaire, L.; Ploman
DNA Cell Biol. 14, 971-981, 1995
A:Title: Developmentally regulated mouse gene NK10 encodes a xZinc Finger Repressor Protein
A:Reference number: I48689
A:Accession: I48689
A:Status: preliminary; translated from GE/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-636 <RES>
A:Cross-references: EMBL:X79828; NID:q506501; PID:q506502

```

Query Match      11.3%  Score 309;  DB 2;  Length 636;
Best Local Similarity 22.9%;  pred. No. 8.8e-15;
Matches 116;  Conservative 50;  Mismatches 140;  Indels 200;  Gaps 19;

Qy  49  EDEDAGEDSMKVKDEYSRD--ENIMKP--EPMGDAAE--SEM-----PY----- 88
      |||||  - - - - - : : :  - - - - - : : :  - - - - -
Db  153  EQNKVGEDSSLDLVPQDLDSSSIRPSDCKTFGNLHNSLVTQSNILAKKPYKCDK 212

Qy  88  ---SVARYESDYESTIKLRRHVPYDNSR---PTSGKMN---CDVGLGSCISFNVLVMVHRKSH 139
      - - - - - : : :  - - - - - : : :  - - - - - : : :  - - - - -
Db  -213  CRKSFIFHRSSLNKKHEIKHGQGYVNGTPOGAOSGRKHHCEACDGGKTFWRLTQLEPHORIH 272

```

Qy	140	TGERPFCNOCGASFTQKGNLRLRIKLHTGEKPKCHLCNLYACQRDALTGHLRTHSVSK	199
Db		: : : :	
Db	273	TGEXPFCNVCGRKAFRHSSSLGHENAHTGEKPYQCGLCAFQSGSSLVQHORIHTGK	332
Qy	200	PYKCEFGCRSYKQBSLEEHK-----ERC-----	224
Db	333	PIYCNLCGRSFRHSTLSUTQHEVTHSGEKKPQCKECKCAFSRCSSLVQHERTHTGKPFEC	392
Qy	224	---RAFLQNPLDGLDAASVEARHKAEKMGSERALVLDRLASNVAKKSMPQKFICEKRH	279
Db	393	SICGRAFGQSPL-----YKHWRI-----HKRSKP-----	418
Qy	280	CFDANYNGMYEKENEMMOTRMWDAQINNAISYLGAERPLVLQTPPTSEMPVVIS	339
Db	418	-----YQSNFSL-----AFVP--NTP-----	433
Qy	340	VYPALTADWPACAPOMEKKRILLPEKILPSERGLSPNNSAQD--STDSDSHEDROHL	398
Db	433	-----LPQG-----EGLLTEVKSCHNCDCGKGFDGHTDFSEHORLHA	469
Qy	399	YOQSHVVLPQARNGMPLLKEVPRSFELLKPPICLRDSIKVINKEGVMDVFCDHCIVL	458
Db	470	GENSY-----GSQTLLGQOSLSHPREKP-----YQCNVCGKA	502
Qy	459	FLDYVMFTIHMGCHGFRDPFECNMCG	484
Db	503	KFRSTSFIHHRIHTGEKPYECNECG	528
RESULT	9		
A37107	spermatogenesis protein zfp-35 - mouse		
N:	Alternate names: finger protein Zfp-35		
C:	Species: Mus musculus (house mouse)		
C:	Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 10-Sep-1991		
C:	Accession: A37107; S07667		
R:	Cunliffe, V.; Williams, S.; Trowsdale, J.		
Genomics	8, 331-339, 1990		
A:	Title: Genomic analysis of a mouse zinc finger gene, zfp-35, that is up-		
A:	Reference number: A37107; MUID:91065652		
A:	Accession: A37107		
A:	Status: preliminary		
A:	Molecule type: DNA		
A:	Residues: 1-580 <CUN>		
A:	Cross-references: GB:J04770		
R:	Cunliffe, V.; Koopman, P.; McLaren, A.; Trowsdale, J.		
EMBO	J. 9, 197-205, 1990		
A:	Title: A mouse zinc finger gene which is transiently expressed during sper-		
A:	Reference number: S07667; MUID:90107944		
A:	Accession: S07667		
A:	Molecule type: mRNA		
A:	Residues: 1-580 <CUN2>		
A:	Cross-references: EMBL:X17617; NID:g55473		
C:	Keywords: DNA binding; transcription regulation; zinc finger		

Query Match	11.3%;	Score 307;	DB 2;	Length 580;
Best Local Similarity	22.5%;	Pred. No. 1.1e-14;		
Matches 139; Conservative	75;	Mismatches 204;	Indels 200;	Gaps 29;

Qy	5	QPTVELKSTTEOPLTESPDALNDYSLPKPHEIENVDVSREAPANEDAGEDSMVKVDEY	64
Db	26	EPATBEEMTGKIGTVEESGLEE---DVPH-----DSRGKEPRFGEELNDQMLFRRRQ	77
Qy	65	SRDE-----NIMKP-----EPMGDA-----EESMPY-----	88
Db	78	YNCDQDSFAWSTGLIRHQRTHKNPYECECGKAFRMSSALVHQRIHTGKEKYPCSWC	13
Qy	88	--SYAREYSDYESIKLER-----HVPYDNSRPTSGKMNCVGLSCITFNVLMVHRKSHTG	14
Db	138	IKSFSSRS--SDL--IKHORVHTGKXPY-----KCDGCKAFSOSDLMITHORIHTG	18

QY 224 RAFLQNPDLGDAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRCFPA 283
Db 170 -----DSAN----- 174
QY 284 NYNPGYMEKENEMOTRMMDQAINNAISYLGAEAFRLVQTPPTAPTSMPVVISVYPI 343
Db 174 -----YEKE-DMTSHVMDQAINNAISYLGAEAFRLVQTPPG-SSEVPVVISMWQL 224
QY 344 ALTRADMP-----MGAPQEMKRIILLPEKILPSERGLSPNNSAQDSTDSDNHEDRQH-- 398
Db 225 HKPPSDGPPRSHNSAQDAVDNLLLSKAKSVSREASPSNCQDSTDFTESNAEQRSGL 284
QY 398 LYQOSHVLVPOARNGMPLKVEPRFELLKPPICLRDSIKVINKEGEVMDVFRDCHV 457
Db 285 IYLTNH-INPHARNGL-ALKEQRAYEVLRAASNSQDAFRVYSTGEOLKYKCEHCRV 342
QY 458 FLFDYVMTIHMCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 503
Db 343 LFLDHWYTIHMCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 388
RESULT 5
S22954
finger protein zfp-37 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C:Accession: S22954
R:Burke, P.S.; Wolgemuth, D.J.
Nucleic Acids Res. 20, 2827-2834, 1992
A:Title: Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmental
A:Reference number: S22954; MUID:942310982
A:Accession: S22954
A:Molecule type: mRNA
A:Residues: 1-553 <BUR>
A:Cross-references: EMBL:X64413; NID:g55474; PID:g55475
C:Superfamily: zinc finger protein ZFP-36
C:Keywords: DNA binding

Query Match 12.6%; Score 344; DB 2; Length 553;
Best Local Similarity 23.3%; Pred. No. 2.3e-17;
Matches 114; Conservative 55; Mismatches 129; Indels 192; Gaps 16;
QY 21 ESPDALNDYSLPKPHEIENVDSREAPA-----NEDEAGEDSMKVDEYSRDENI 71
Db 121 KPPTANEHRKSLSHSASDVNDIEPTRKCKDLNPNKLSKRGD-----KNOTSKCKEV 175
QY 72 MKPEPMGDAEE---SEMPYSYAREYSDYESIKLERHVPYDNRPTSGK-MNCDVCGVLS 126
Db 176 CHRSHSHTKEDKIOTGEKRRKSHCRTPSKPE-----KAPGSGKPYECNHCQKVL 223
QY 127 ISFNVLVHKRSHTGERPQCQCGASFTQKGNLLRHIKLHTGKPFKCHLCNACQRRD 186
Db 224 SHKQGLLDHQRHTGKPYECNECGIAFSQKSHLVVHQHTHTGKPYECEQCGKAHGHK 283
QY 187 ALTGHLRTHSVKPYKCFGRSYKORSLEEH-----KERCRAFLQNPDLGD 234
Db 284 ALTDHLRIHTGKPYKCNCEGTFRSHSNLMQHLRSHTGKPYECKGCKSPRYNSSL-- 342
QY 235 AASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRCFDPNPGYMEKE 294
Db 342 -----TEHVRTHTG-----EIPY-----ECNEC----- 360
QY 295 NEMOTRMMDQAINNAISYLGAEAFRLVQTPPTAPTSMPVVISVYPIALTRADMPGA 354
Db 360 -----GKAFK-----YGSLLTK----- 372
QY 355 POEMKKRILLPEKILPSERGLSPNNSAQDSTDSDNHEDRQHLVQOSHVLVPOARNGM 414
Db 372 -----HMRIHTGK--PFE-----CNECGTFSKSHLVHQHTH-- 405
QY 415 LLKVPKRSFELLKPPICLRDSIKVINKEGEVMDVFRDCHVFLDYVMTIHMCHGCF 474
Db 111 ----- 111

Db 405 -TKKP-----YKDECGKAPGHSSSLTYHMRHTTG 434
QY 475 RDPFECNMG 484
Db 435 DCPFECNMG 444
RESULT 6
S10245
finger protein, testis - mouse
C:Species: Mus musculus (house mouse)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 10-Sep-1997
C:Accession: S10245
R:Neikii, D.; Dudley, K.; Cunningham, P.; Akhavan, M.
Nucleic Acids Res. 18, 3655, 1990
A:Title: Cloning and sequencing of a zinc finger cDNA expressed in mouse testis.
A:Reference number: S10245; MUID:90301500
A:Accession: S10245
A:Molecule type: mRNA
A:Residues: 1-411 <NEL>
A:Cross-references: EMBL:X52533; NID:g53456; PID:g53457
C:Keywords: DNA binding; zinc finger
Query Match 11.7%; Score 319; DB 2; Length 411;
Best Local Similarity 24.6%; Pred. No. 9.8e-16;
Matches 95; Conservative 41; Mismatches 88; Indels 162; Gaps 12;
QY 112 PTSGK-MNCDVCGVLSISFNVLVHKRSHTGERPFCQCGASFTQKGNLLRHIKLHTGE 170
Db 66 PGGKPYECNHCQGVLSHKQGLLDHQRHTHTGKPYECNECGIAFSQKSHLVHQHTTG 125
QY 171 KPFKCHLCNACQRRDALTGHLRTHSVKPYKCFGRSYKORSLEEH----- 220
Db 126 KPYECEQCGKAHGHKHALTDHLRIHTGKPYKCNCEGTFRSHSNLMQHLRSHTGKPYE 185
QY 220 -KERCRAFLQNPDLGDAASVEARHIKAEMGSEALVLDRLASNVAKRKSSMPQKFIGEKR 278
Db 186 CKEGCKSPRYNSSL-----TEHVRTHTG-----EIPY-----ECN 215
QY 279 HCFDANYNPGYMEKENEMOTRMMDQAINNAISYLGAEAFRLVQTPPTAPTSMPVIS 338
Db 216 EC-----GKAFK----- 223
QY 339 SVYPIALTRADMPGAPOEMKKRILLPEKILPSERGLSPNNSAQDSTDSDNHEDRQHL 398
Db 223 -YGSLLTK-----HMRIHTGK--PFE-----CNECGKTF 249
QY 399 YQOSHVLVPOARNGMPLKVEPRFELLKPPICLRDSIKVINKEGEVMDVFRDCHV 458
Db 250 SKKSHLVHQHTH--TKEP-----YKDECGKA 276
QY 459 FLDYVMTIHMCHGFRDPFECNMG 484
Db 277 FGHSSSLTYHMRHTGDCPFECNMG 302
RESULT 7
S00647
finger protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
C:Accession: S00647
R:Ruiz i Altaba, A.; Perry-O'Keefe, H.; Melton, D.A.
EMBO J. 6, 3065-3070, 1987
A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
A:Reference number: S00647; MUID:88082679
A:Accession: S00647
A:Molecule type: mRNA
A:Residues: 1-1350 <RUI>
A:Cross-references: EMBL:X06021
A:Note: It is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the Initi
C:Genetics:

C:Accession: I59572
R:Georgopoulos, K.; Moore, D.D.; Derfler, B.
Science 238, 808-812, 1992
A:Title: Ikaros, an early lymphoid-specific transcription factor and a putative mediator of B cell development
A:Reference number: I59572; MUID:93086267
A:Accession: I59572
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-431 <RES>
A:Cross-references: GB:I03547; NID:g198286; PID:g198287
C:Genetics:
A:Gene: Ikaros

Best Local Similarity	51.3%;	Pred. No. 9.2e-76;	
Matches 245; Conservative	64;	Mismatches	93;
		Indels	76;
		Gaps	15;

:: : | : | : | : | :
 : | : | : | : | :
 : | : | : | : | :

Country	Year	Value
Algeria	1990	1.00
Algeria	1991	1.00
Algeria	1992	1.00
Algeria	1993	1.00
Algeria	1994	1.00
Algeria	1995	1.00
Algeria	1996	1.00
Algeria	1997	1.00
Algeria	1998	1.00
Algeria	1999	1.00
Algeria	2000	1.00
Algeria	2001	1.00
Algeria	2002	1.00
Algeria	2003	1.00
Algeria	2004	1.00
Algeria	2005	1.00
Algeria	2006	1.00
Algeria	2007	1.00
Algeria	2008	1.00
Algeria	2009	1.00
Algeria	2010	1.00
Algeria	2011	1.00
Algeria	2012	1.00
Algeria	2013	1.00
Algeria	2014	1.00
Algeria	2015	1.00
Algeria	2016	1.00
Algeria	2017	1.00
Algeria	2018	1.00
Algeria	2019	1.00
Algeria	2020	1.00
Algeria	2021	1.00
Algeria	2022	1.00
Algeria	2023	1.00
Algeria	2024	1.00
Algeria	2025	1.00
Algeria	2026	1.00
Algeria	2027	1.00
Algeria	2028	1.00
Algeria	2029	1.00
Algeria	2030	1.00
Algeria	2031	1.00
Algeria	2032	1.00
Algeria	2033	1.00
Algeria	2034	1.00
Algeria	2035	1.00
Algeria	2036	1.00
Algeria	2037	1.00
Algeria	2038	1.00
Algeria	2039	1.00
Algeria	2040	1.00
Algeria	2041	1.00
Algeria	2042	1.00
Algeria	2043	1.00
Algeria	2044	1.00
Algeria	2045	1.00
Algeria	2046	1.00
Algeria	2047	1.00
Algeria	2048	1.00
Algeria	2049	1.00
Algeria	2050	1.00
Algeria	2051	1.00
Algeria	2052	1.00
Algeria	2053	1.00
Algeria	2054	1.00
Algeria	2055	1.00
Algeria	2056	1.00
Algeria	2057	1.00
Algeria	2058	1.00
Algeria	2059	1.00
Algeria	2060	1.00
Algeria	2061	1.00
Algeria	2062	1.00
Algeria	2063	1.00
Algeria	2064	1.00
Algeria	2065	1.00
Algeria	2066	1.00
Algeria	2067	1.00
Algeria	2068	1.00
Algeria	2069	1.00
Algeria	2070	1.00
Algeria	2071	1.00
Algeria	2072	1.00
Algeria	2073	1.00
Algeria	2074	1.00
Algeria	2075	1.00
Algeria	2076	1.00
Algeria	2077	1.00
Algeria	2078	1.00
Algeria	2079	1.00
Algeria	2080	1.00
Algeria	2081	1.00
Algeria	2082	1.00
Algeria	2083	1.00
Algeria	2084	1.00
Algeria	2085	1.00
Algeria	2086	1.00
Algeria	2087	1.00
Algeria	2088	1.00
Algeria	2089	1.00
Algeria	2090	1.00
Algeria	2091	1.00
Algeria	2092	1.00
Algeria	2093	1.00
Algeria	2094	1.00
Algeria	2095	1.00
Algeria	2096	1.00
Algeria	2097	1.00
Algeria	2098	1.00
Algeria	2099	1.00
Algeria	2100	1.00
Algeria	2101	1.00
Algeria	2102	

```

164 IKLHTGEKPFKCHLCNYACQRRDALTGHLRTHSVEKPYKCFGRSYKQRSSLEEHKERC 223
QY

```

QY 224 RAFLQNPD LGDAASV --EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPQKFIG 275

DD HNI LESMGLEFGVCFV IAKEIINHNEAEDEUCRIGAEKSLVLDRLASNVAKRKSSMPQKFLG 196

Db 197 DK--CLSDMPYDSA-NYEKE-DMMTSHVMDQAINNAINYLGAESLRPLVQTPPG-SSEVV 251

Db	252	PVISSMYQLHKPPSDGPPRSNHSQAQDAVDNLLLSKAKSVSSEREASPSNSCODSTD	TES	311

213 NAEF08CCTV1E7M1 - TAPUS09CCTV1E7M1

||:::|| :| :| |||: || |::|: || ::|: || :

QY 443 VERCDHCCHVLEDDI VMF I I H M - - - GCHGF K D F F E C N M C G I K S H D R I E F S S H I A R G E H R 303

2
RESULTS

C:Species: Mus musculus (house mouse)

R;Hahn, K.; Ernst, P.; Lo, K.; Kim, G.S.; Turck, C.; Smale, S.T.
Mol. Cell Biol. 14: 7111-7123 1994

A:Title: The lymphoid transcription factor Lyf-1 is encoded by specific, alternatively spliced transcripts
A:Reference number: A56229; MUID:95021239

Molecule type: mRNA
Residues: 1-427

Cross-references: GB:
Keywords: alternative

Best Local Similarity 50.9%; Pred. No. 1.3e-75;

—

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:07:40 ; Search time 18.18 Seconds
(without alignments)
1117.338 Million cell updates/sec

Title: US-09-019-348-2
Perfect score: 2728
Sequence: 1 MEDIQTVELKSTEEQLPT.....HDIYFSSHIARGEHRAMLK 507

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_60.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1286	47.1	518	2 A56355	DNA-binding protei
2	1156	42.4	431	2 I59572	Ikaros DNA binding
3	1154	42.3	427	2 B56229	Lymphoid transcrip
4	664	24.3	392	2 A56229	Lymphoid transcrip
5	344	12.6	553	2 S22954	finger protein zip
6	319	11.7	411	2 S10245	finger protein, te
7	316	11.6	1350	2 S06647	gene NK10 protein
8	309	11.3	636	2 I48689	spermatogenesis pr
9	307	11.3	580	2 A37107	zinc finger protei
10	298.5	10.9	543	2 B34612	zinc finger protei
11	298	10.9	439	2 S06536	finger protein (cl
12	297.5	10.9	595	2 G02075	transcription repr
13	296.5	10.9	510	2 A36901	chorion transcrip
14	295.5	10.8	675	2 S51037	zinc-finger protei
15	295.5	10.8	803	2 S26823	finger protein ZNF
16	295	10.8	693	2 I37570	zinc finger protei
17	293.5	10.8	485	2 A40751	finger protein ZNF
18	293.5	10.8	393	2 JN0533	finger protein pML
19	293.5	10.8	614	2 JH0500	zinc finger protei
20	293	10.7	295	2 A46017	zinc finger protei
21	292.5	10.7	1191	2 S35305	zinc finger protei
22	292	10.7	378	2 S00549	developmental cont
23	291	10.7	378	2 S33994	finger protein ZNF
24	289.5	10.6	491	2 S00753	finger protein HF
25	288.5	10.6	519	2 A38073	transcription acti
26	288	10.6	545	2 I49636	DNA-binding protei
27	286.5	10.5	651	2 B32891	finger protein 2,
28	286.5	10.5	292	2 S43826	finger protein OZF
29	286	10.5	428	2 S03677	finger protein (cl
30	284.5	10.4	247	2 S47070	finger protein HZF
31	283.5	10.4	591	2 S65088	finger protein XFO
32	283.5	10.4	427	2 A35659	krueppel-related p
33	282	10.3	474	2 I54338	zinc finger protei
34	281	10.3	247	2 S06553	finger protein (cl
35	279	10.2	732	2 S47073	finger protein HZF
36	277.5	10.2	671	2 JE0288	krueppel-type zinc
37	277.5	10.2	542	2 A54661	zinc finger protei
38	277	10.2	428	2 A32891	finger protein 1,
39	277	10.2	469	2 I38600	zinc finger protei

finger protein kox
finger protein (cl
finger protein XFO
zinc finger protei
213 protein - mous
renal transcriptio

ALIGNMENTS

RESULT 1

A56355

DNA-binding protein Ikaros form 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995

C:Accession: A56355

R:Molnar, A.; Georgopoulos, K.

Mol Cell Biol. 14, 8292-8303, 1994

A:Title: The Ikaros gene encodes a family of functionally diverse zinc finger DNA-bin

A:Reference number: A56355; MUID:95059058

A:Accession: A56355

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-518 <MOL>

C:Keywords: alternative splicing; DNA binding

Query Match 47.1%; Score 1286; DB 2; Length 518;

Best Local Similarity 53.1%; Pred. No. 55e-85;

Matches 276; Conservative 69; Mismatches 129; Indels 46; Gaps 17;

Qy 12 STEEQPLTSPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVDEYSDRDENI 71

Db 13 SGKESPPVSDTPDE-GDEPMVPVEDLSTTSG--AQQNSKSDRGWAS-NVKVETQSDENG 68

Qy 72 MKPEPMG-----DAESEMPPSYAREYSDYESIKLERHVPYDINSRPTSGKMNCDV 121

Db 69 RACEMNGEECAEDLUMLDASGERKMGSHRDQGSALS-----GVGGIRLPNGKLKCDI 121

Qy 122 CGLSCIGFNVLVHVKRSHTERPFCQNOCCASFTQKGNLLRHKLHTGERPKFCHLCNYA 181

Db 122 CGIVCIGFNVLVHVKRSHTERPFCQNOCCASFTQKGNLLRHKLHTGERPKFCHLCNYA 181

Qy 182 QQRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEKERCRFLQN---PDLDGAASV 238

Db 182 CRRDALTGHLRTHSVKPKYKCEFCGRSYKORSLEEKERCHNYLESMLPGMPYVIKE 241

Qy 239 EARHIK-----AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKHCF-DANYNPGYME 292

Db 242 ETNENMAEDLCKTGAERSLVLDRLASNVAKRKSSMPQKFIGDK--CLSDMPYDSA-NYE 298

Qy 293 KENEMOTRMMDQAINNAISYLGAERPLVQTPPTSEMPVVISVYPITALTRADMP- 352

Db 299 KE-DMTSHVMDQAINNAISYLGAERPLVQTPPG-SSEVPVISVYQHLKPPSGPP 356

Qy 352 ---MGAPQEMKEKRILLKIPLSERGLSPNNSAQDSTDSDSNHEDRQH---LYQQSHVYL 406

Db 357 RSNHSAQDAVDNLLLSKAKSVSSEREASPSNQCQSDTDTESNAEQRSGLIYLTNH-IN 415

Qy 407 POARNGHPLLEKVPDRSELKPPICLDSDIKVINKGEVMDVFCRCHVFLDYVMT 466

Db 416 PHARNGI-ALKEQORAYEVLRAASENSQDAFRVYSTGEQLKYKCEHCRVFLDHVMT 474

Qy 467 IHM---GCHGFRDPFECNMGYRSHDRYFSSHIARGEHR 503

Db 475 IHMCHGCHGFRDPFECNMGYRSHDRYFSSHIARGEHR 514

RESULT 2

I59572

Ikaros DNA binding protein - mouse

C:Species: Mus musculus (house mouse)

THIS PAGE BLANK (USPTO)

[illegible]

Search completed: November 6, 1999, 09:41:35
Job time: 67 sec

[illegible]

```

RESULT 14
US-08-718-661-2
: Sequence 2, Application US/08718661
: Patent No. 5876972
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
: : tumor suppressor proteins and methods for their isolation
: :
: NUMBER OF SEQUENCES: 15
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release. #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/718,661
: FILING DATE:
: CLASSIFICATION: 530
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 667 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-718-661-2

```

```

Query Match      8.2%; Score 224.5; DB 2; Length 667;
Best Local Similarity 21.7%; Pred. No. 6e-14;
Matches 89; Conservative 52; Mismatches 166; Indels 103; Gaps 14;

Qy 119 CDVGLSGISFNVLWVKRSHSTGTERPQCN--QCGASPTQGNLLRHLTKLHTGKPFKCH 176
      | | | | : | | : | | | | : | | : | | | : | | : | | : | |
Db 6 CQKCGKSVTLTKETTHYSHSRERPKCSKAECCGKAFVSKYKLMRHMATHSPQKHQCT 65

Qy 177 LCNTACQRDALTGHLRTHSVK-PYKCEFGCRSYKQRSSLEEHKRCRAFLQNPDLGDA 235
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 HCEFTNRKDKHLKNHLOTHDPNKISYACDCCGKRYHTMLGKYKH-----LALHSASNGD- 120

```

```

236 ASVEARIHKAEMGERALVLDRIASVNAKRKSSMP----- 271
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
120 --LTCGVCTLELGSTEVL-LDHLKSHAEKANAQPREKKYQCDHCDCRFYTRKDRVRLHV 176
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
271 -----QKFIGEKRC-----FDANYN--PGMYVE 292
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
177 VHTGCKDFLCQCAQRF-GRKDLHTRHTKTHSQELMQENMGDYGQSNFOLIARSTSQ 235
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
293 KENEMGOTRMDQAINNAIS-----YLGAEAFRLVOTPTAPPTSEMVPV--- 337
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
236 IKVDPMPFPFOLGAAPENGLDGLPPEVHGVLAAPEAPQMPPELPLEPLEPLEPM 295
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
338 SSVYPIALTRADPMGAQPEMEKRIILP-----EKLPSF--RGLSPNNSAODSTDTSN 391
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
296 QSLPLEPLEPMQLEPMQPLEPMQPLEPMQPLEPMQPLEPMQPLEPMQPMQPMQPM 355
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
392 HEDRQHLYQQSHVVLQFARNGMPLLKVEPRSFEL-----LKPPPICLRD 435
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
356 QPMQMLPMQPMPLPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPMQPM 405
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

RESULT 15
US-08-040-548-1
: Sequence 1, Application US/08040548
: Patent No. 5763209
: GENERAL INFORMATION:
: APPLICANT: Sukhame, Vikas P.
: TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE
: TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS
: NUMBER OF SEQUENCES: 67
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, white & Durkee
: STREET: 321 No. 5763209th Clark Street, Suite 800
: CITY: Chicago
: STATE: Illinois
: COUNTRY: U.S.A.
: ZIP: 60610
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/040,548
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Coughlin, Daniel F.
: REGISTRATION NUMBER: 36,111
: REFERENCE/DOCKET NUMBER: arcd067
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 744-0090
: TELEFAX: (312) 245-4961
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 533 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-040-548-1

```

```
Query Match      8.0%; Score 218.5; DB 2; Length 533;
Best Local Similarity 26.4%; Pred. No. 1.7e-13;
Matches 71; Conservative 23; Mismatches 80; Indels 95; Gaps 10;
```

Qy 142 ERPQC--NOCGAFSTQGNLLRIKILHTEKKPKFCHLCNYACORRDALGLRTHSVEK 199
||| | | ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 333 ERPYACVSCDRFRFSDELTRIRHTQGKPPQCRICMRNSRSDLHTHTIHTHGKEK 392

Qv -200 PYKEFCGRSYKQRSSLEEHEKRCRAFLQNPDLGDAASVEARHIKAEMGERALLVDRLA 259

; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-074-967-2

Query Match 9.7%; Score 265.5; DB 1; Length 706;
Best Local Similarity 27.8%; Pred. No. 5.3e-18;
Matches 70; Conservative 39; Mismatches 88; Indels 55; Gaps 9;

QY 5 OPTVELKSTEQPLTESPDALNDYSLPKPHEIENVDSRE---APANEDE----- 52
Db 415 QPPMEPNLDLQS-PTKLSASGEDSTIPQASRLNNIVNRSMTGSPRSSSESHSPLYMHPP 473

QY 52 ---DAGDSMKVDEYSDRDENIMKP---EPMGDAESEMYPYAREYSD----- 96
Db 474 KCTSCGSQS---PQAEMLHTAGTFAEEMGETQS-----EYSDSCENGAFCC 520

QY 96 -----YESIKLERHVPYDNS-RPTSGKMNCDVCGLSCLISFNVLVHKRSHGTGERPFC 147
Db 521 NECDCRFSEASLKRHTLTQTHSDKP---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 148 NOCGASFTQKGNLRLHKLTGKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCG 207
Db 577 NICGAQFNRPANLKTHTRIHSGKPKYCETCGARFVQVAHLRAHVLHTGKPYPCICG 636

QY 208 RSYKQSSLEEH 219
Db 637 TRFRHLQTLKSH 648

RESULT 10
US-08-553-541B-2
; Sequence 2, Application US/08553541B
; Patent No. 5882858
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Chaganti, Raju S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
; TITLE OF INVENTION: bcl-6
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,541B
; FILING DATE: May 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-553-541B-2

Query Match 9.7%; Score 265.5; DB 2; Length 706;
Best Local Similarity 27.8%; Pred. No. 5.3e-18;
Matches 70; Conservative 39; Mismatches 88; Indels 55; Gaps 9;

QY 5 OPTVELKSTEQPLTESPDALNDYSLPKPHEIENVDSRE---APANEDE----- 52
Db 415 QPPMEPNLDLQS-PTKLSASGEDSTIPQASRLNNIVNRSMTGSPRSSSESHSPLYMHPP 473

QY 52 ---DAGDSMKVDEYSDRDENIMKP---EPMGDAESEMYPYAREYSD----- 96
Db 474 KCTSCGSQS---PQAEMLHTAGTFAEEMGETQS-----EYSDSCENGAFCC 520

QY 96 -----YESIKLERHVPYDNS-RPTSGKMNCDVCGLSCLISFNVLVHKRSHGTGERPFC 147
Db 521 NECDCRFSEASLKRHTLTQTHSDKP---YKCDRCQASFRYKGNLASHKTVHTGKPYRC 576

QY 148 NOCGASFTQKGNLRLHKLTGKPFKCHLCNYACORRDALTGHLRTHSVKPKYKCEFCG 207
Db 577 NICGAQFNRPANLKTHTRIHSGKPKYCETCGARFVQVAHLRAHVLHTGKPYPCICG 636

QY 208 RSYKQSSLEEH 219
Db 637 TRFRHLQTLKSH 648

RESULT 11
PCT-US94-06669-2
; Sequence 2, Application PC/TUS9406669
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Chaganti, R.S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
; TITLE OF INVENTION: bcl-6
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06669
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/074,967
; FILING DATE: 09-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43771-A-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear


```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 641127
; US-08-933-750C-17

Query Match 10.1%; Score 275.5; DB 2; Length 488;
Best Local Similarity 25.3%; Pred. No. 3e-19;
Matches 80; Conservative 47; Mismatches 122; Indels 67; Gaps 11;

QY 13 TEQPLPTEPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSMKVKDEYSDRDENIM 72
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 174 TEKPCDHNCGKILNTS-PDGHPIYKIHAE-KQYEGSCGKN-----FSOSSELL 224
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 73 KPMPGDAESEMPPY-----SYARYSDYESIKLERHVPYDINSRPTSGKMNCDCVGLSCISF 129
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 225 HQRD-----HTEEPYKCECGKGFTRSSLLIHQAVHTDEK-----PYKCDKCGKGFTRS 275
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 130 NVLMVHRKSHGTGPFQCNOCGASFTQGNLLRHKIHTGKPFKCHLCNYACORRDALT 189
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 276 SLLIHAVHTGKPKYKCDKCGKGFSSQSKLHIHQVHTGKPYECCECGMSFSQSRNLH 335
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 190 GHLRTHSVKPYKCEFCGRSKYKORSSLEEHKERC-----RAFLQNPDLGDA 235
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 336 IHQVHTGERPKYKCGCGKGFSSQSNL--HIHRCIHTGKPYOCYECGKGFSSQSDL--- 391
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 236 ASVEARIHKAEMG-----SERALVLDRL-----ASNVAKRKSSMPQK 272
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 391 ----RIHLRVHTGKPYKHCCKGKGFSSQSKLLIHQVHTGKPYECCKGKGFSSQSNL 446
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 273 FIGEKRCFCFDANYNGP 288
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 447 HIHQVHKRDPRAHPG 462
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 6
US-08-820-170A-10
; Sequence 10, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States

; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-820-170A-10

Query Match 10.0%; Score 273; DB 2; Length 711;
Best Local Similarity 30.0%; Pred. No. 9.6e-19;
Matches 62; Conservative 30; Mismatches 63; Indels 52; Gaps 4;

QY 119 CDVCGSCISFNVLMVHRKSHGTGERPFQCNOCGASFTQGNLLRHKIHTGKPFKCHLC 178
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 477 CHNCGKSFISQOLDIHRHTGKPYECDCGKTFTQKSHLNIHQIHTGERHHVCSSEC 536
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 179 NYACQRRDALTGHLRTHSVKPYKCEFCGRSKYKORSSLEEHKE----- 222
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 537 GKAFNQKSLMSHQRIHTGKPYKCECCKAFTSKSQFKEHQRIHTGKPYVCTCGRAF 596
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 222 -----RC-RAFLQNPDLGDAASVEARIHKAEMGSEALVLDRL 258
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 597 NGRSNEFKHQIHTRERPFVCKGKAFVKSEL-----ITHQTHMGKPYECLD--- 648
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
QY 259 ASNVAKRKSSMPQKFTGKRCRHFCDANY 285
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |
Db 648 ---CGKSFSSKPKLVHQRIHTGERPY 671
| | | | | : : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 7
US-08-933-750C-14
; Sequence 14, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
```

PCT-US95-09345-13

Query Match 44.3%; Score 1208; DB 3; Length 470;
Best Local Similarity 58.3%; Pred. No. 2.2e-112;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;

QY 111 RTSKMNCDCVGLSCISFNVLVHKSHTGERPPFCNOCGASFTQKGNLLRHKLHGTGE 170
DB 60 RUPNGLKCDIGICIGPNVLVHKSHTGERPPFCNOCGASFTQKGNLLRHKLHGTGE 119
QY 171 KPFKCHLCNACORDALHTSHVEKPKYCEFCGSRYSKORSLSEEHKERCRAFLONP 230
DB 120 KPFKCHLCNACRRDALHTSHVSKPKHCGCGSRYSKORSLSEEHKERCCHNYLES 179
QY 231 DL-GDAASY---EARRHK-----AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKHCF 281
DB 180 GLPGXXXPVIKETXHXEMAEADLCCKIGXERSLVLDRLASNVAKRKSSMPQKFLGDK-XLS 238
QY 282 DANYNPGWYKENEMQMTRMDQAINNAISYLGAEAPRLVQTPPATSEWVPVIVSY 341
DB 239 DXPYDSA-XYEREXXMXSHVMDXAINNAISYLGAEAPRLVQTPPG-XSEYVVPVISP 296
QY 342 PIALTRADMPGAPOMEKKR-----ILLPEKILPSERGLSPNNSAODSTDSDSNHE 393
DB 297 QL---HXXSXCPXSNHSAQXAVXXLLLSKAKXXVXSERASPSNCSQDSTDSTESN 353
QY 394 DRQH--LYQSHVLPQARNGMPLKEVPRSFELKPPPICULRDSIKVINKEGEYMDVFR 451
DB 354 EQRSGLIYLTNIX--XXAXXXXLKEEXRAYXXLRAASENSQDAXRVVSTSGEQQKVK 411
QY 452 CDHCHVFLDYWFYIHM---CCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 503
DB 412 CEHCRVFLDHYWYIHMXXGXCHGFRDPFCNMGCGYRSHDRYEFSSHIARGEHR 466

RESULT 4

US-08-933-750C-4
; Sequence 4, Application US/089333750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933-750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 9337
; US-08-933-750C-4

Query Match 10.6%; Score 289.5; DB 2; Length 338;
Best Local Similarity 26.7%; Pred. No. 6.6e-21;
Matches 82; Conservative 49; Mismatches 109; Indels 67; Gaps 12;

QY 1 MEDIQTV--ELKSTEQPLPTE-----SPDAL--NDYSLPKPHEIENVDSREAP 46
DB 60 VOEVQDITLVKQMSAQEKDLPQKKHFDNRESQANSGLDITNOVSLQK-----IDNPESQ 113
QY 47 ANEDDAGEDSMKV-----KDEYSRDENI---MKPEPMGDABESEMPSYAREYSDE 97
DB 114 AN--SGALDTNQVLLHKKIPPRKRLKRDQVSKMKHNSRVKIHQKSCERQKAKENGCR 170
QY 98 SI--KLERHVPYDNRSTSGKMCNDCVGLSCISFNVLVHKSHTGERPPFCNOCGASFT 155
DB 171 KTFSRSTKQITFIIRHKGQVRCSECGKIFRNPYFSVHKKIHTGERPYVCQCGKGFV 230
QY 156 QKGNLLRHKLHKTGEKPKFCHLCNACORRALTGHLRHTSHVEKPKYCEFCGSRYSKORSS 215
DB 231 QSSSLTQHQVHSGERPEFCEQCGTNDRSASISQHLRHTGTCARPYKQDCGCKAFRQSSH 290
QY 216 LEEHK-----ERC-RAFLQNPDLGDAASVEARHKAEMGSEALVLDRLASNVA 263
DB 291 LIRHQHTHTGERPYACNKGKGAFTOSSHL-----IGHQR-----THNRT 329
QY 264 KRKSSMP 270
DB 330 KRKKKQP 336

RESULT 5

US-08-933-750C-17
; Sequence 17, Application US/089333750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

QY 366 PEKILPSERGLSPNNSAQDSTDDTDSDNEDRQH--LYQOQSHVVLPOARNGMPLLEKEVPRSF 423
Db 424 KAKSVSSREASPSNCSODSTDTESNAEORSGLIYLTNH-INPHARNGL-ALKEEQRAY 481
QY 424 ELLKPPPTCLRLDSIKVINKKEGVMDVRCDBCHVFLFDYVNFTHM--GCHGFRDPFEC 480
Db 482 EVLRAAENSODAFRVYSTGEQLKVYKCEHCRVFLFDHVMYTHMCHGCHGFRDPFEC 541
QY 481 NMCYGRSHDRYEFSSHIARGHR 503
Db 542 NMCYGRSHDRYEFSSHIARGHR 564

RESULT 2
US-08-465-590-153
; Sequence 153, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,590
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA: US 08/238,212
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal
US-08-465-590-153

Query Match 44.3%; Score 1208; DB 2; Length 470;
Best Local Similarity 58.3%; Pred. No. 2.2e-112;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;
QY 111 RPTSGKMNCDVGLSCISFNVLVWVKHSHTGERPFCNQCASFQGNLLRHLKLTGE 170
Db 60 RLPNGKLKDCIGXCIGPNVLVWVKHSHTGERPFCNQCASFQGNLLRHLKLSGE 119
QY 171 KPFKCHLCNACQRRDALTGHLRTHSVKPYKCEFCGRSYKQRSSLEEHKRCRAFLQNP 230

Db 120 KPFKCHLCNACQRRDALTGHLRTHSVGPHKCGYCGRSYKORXSLEEHKRCNHYLES 179
QY 231 DL-GDAASY---EARHIK-----AEMGSRALVLDRLASNAVKRKSMPQKFTGKRRHCF 281
Db 180 GLPGXXXPKVKEETXHXEMAEDLCKIGXERSLVLDRLASNAVKRKSMPQKFLGDK-XLS 238
QY 282 DANYNPGYMYKENEMQMOTRMDQAINNAISYLGAFAFRPLVOTPPAPTSEMVVPVTSVY 341
Db 239 DXPYDSA-XYEKEXMMXSHVMDXAINNAIYLGAESLRLVOTPPG-XSEVVPVISP 296
QY 342 PIALTRADMPGAPQEMKKR-----ILLPEKILPSERGLSPNNSAQDSTDDTDSDNHE 393
Db 297 QL---HXXSXGXPRSNHSAQDXAVYXXLLLSLAKAXVXSERASPSNCSODSTDTESN 353
QY 394 DRQH--LYQOQSHVVLPOARNGMPLLEKEVPRSFELLKPPPTCLRLDSIKVINKKEGVMDV 451
Db 354 EQRSGLIYLTNHIX--XXAXXXXLKEEXRAYXXLRAAENSQDAXRVYSTGEQXKVYK 411
QY 452 CDHCHVFLFDYVNFTHM--GCHGFRDPFECNMGYGRSHDRYEFSSHIARGHR 503
Db 412 CEHCRVFLFDHVMYTHMXXXGCHGFRDPFECNMGYGRSHDRYEFSSHIARGHR 466

RESULT 3
PCT-US95-09345-13
; Sequence 13, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.
OM protein - protein search, using sw model
Run on: November 6, 1999, 09:40:28 : Search time 15.19 Seconds
(without alignments)
381.455 Million cell updates/sec

Title: US-09-019-348-2
Perfect score: 2728
Sequence: 1 MEDIQPVELKSTEEQPLPT.....HDIYEFSSHARGHEHRAMLK 507

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/PCTUS9_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1279	46.9	568	3	PCT-US93-08743-5
2	1208	44.3	470	2	US-08-465-590-153
3	1208	44.3	470	3	PCT-US95-09345-13
4	289.5	10.6	338	2	US-08-933-750C-4
5	275.5	10.1	488	2	US-08-933-750C-17
6	273	10.0	711	2	US-08-820-170A-10
7	272.5	10.0	368	2	US-08-933-750C-14
8	266	9.8	727	3	PCT-US95-08429-9
9	265.5	9.7	706	1	US-08-074-967-2
10	265.5	9.7	706	2	US-08-553-541B-2
11	265.5	9.7	706	3	PCT-US94-06669-2
12	262.5	9.6	728	3	PCT-US95-08429-5
13	237	8.6	765	2	US-08-398-590A-40
14	224.5	8.2	667	2	US-08-718-661-2
15	218.5	8.0	533	2	US-08-040-548-1
16	218.5	8.0	533	2	US-08-466-344-1
17	218.5	8.0	496	2	US-08-224-482-2
18	215.5	7.9	543	2	US-08-224-482-4
19	209	7.7	387	2	US-08-224-482-8
20	208	7.6	84	2	US-08-459-568-62
21	208	7.6	84	2	US-08-399-411-62
22	206.5	7.6	429	1	US-08-234-783-4
23	206.5	7.6	429	1	US-08-456-907-4
24	206.5	7.6	456	2	US-08-040-548-2
25	206.5	7.6	456	2	US-08-466-344-2
26	206.5	7.6	455	2	US-08-224-482-6
27	206.5	7.6	429	3	PCT-US95-05523-4
28	206	7.6	89	2	US-08-040-548-8
29	206	7.6	89	2	US-08-466-344-8
30	205	7.5	547	2	US-08-340-203A-3
31	205	7.5	547	2	US-08-452-567-3
32	205	7.5	547	2	US-08-452-427-3
33	203.5	7.5	345	1	US-08-102-942A-2
34	199	7.3	675	1	US-08-317-522A-9
35	199	7.3	675	1	US-08-439-818A-9
36	199	7.3	675	2	US-08-751-965-9
37	199	7.3	675	2	US-08-738-975-9
38	199	7.3	675	2	US-08-728-626-9
39	193.5	7.1	449	1	US-08-102-942A-6

40 192.5 7.1 449 1 US-08-102-942A-4
41 190 7.0 109 2 US-08-224-482-11
42 183.5 6.7 470 2 US-08-946-241B-2
43 183.5 6.7 479 2 US-08-946-241B-9
44 182 6.7 60 2 US-08-117-952-786
45 179.5 6.6 1706 2 US-08-459-568-2

ALIGNMENTS

RESULT 1
PCT-US93-08743-5
; Sequence 5, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US93-08743-5

Query Match 46.98; Score 1279; DB 3; Length 568;
Best Local Similarity 49.48; Pred. No. 2.5e-119;
Matches 278; Conservative 72; Mismatches 131; Indels 82; Gaps 18;
Qy 12 STEQPLPTESPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSMKVDEYSDRDENI 71
Db 13 SGKSPVPVSDTPDE-GDEMPVPEDLSTTSG--AQQNSKSDRGMAS-NVKVETOSDEENG 68
Qy 72 MKPEPMG-----DAESEMPTYSAREYSDYES----- 99
Db 69 RACEMNGEECAEDLRMLDASGEKMGSHRDQGSALSQVGGIRLPNGKXKDCIGVICIG 128
Qy 99 --IKLERHVPYDNR-----TSKMNCDVGLSCISFNVLVHKRS 138
Db 129 PNVLMVHKRSHTGRPFQCNQCSALSQVGGIRLPNGKXKDCIGVICIGPNVLVHKRS 188
Qy 139 HTGRPFQCNQCGASFQKGNLLRHKLHTGEKPFKCHLCNYACORRDLTGHLRTHSVE 198
Db 189 HTGRPFQCNQCGASFQKGNLLRHKLHSGEKPFKCHLCNYACRRDLTGHLRTHSVG 248
Qy 199 KPYKCEFCGRSYKQSSLEEHKERCRAFLQN---PDLGDAASVEARHK-----AEMGSE 250
Db 249 KPHKCGYGRSYKQSSLEEHKERCCHNYLESMLPGMYPVVIKEETNNHMAEDLCKIGAE 308
Qy 251 RALVLDRLASNVARKSMPOKFIGEKHCF-DANYAPGYMYEKENEMQOTRMQDAINN 309
Db 309 RSLVLDRLASNVARKSMPOKFLGDK--CLSDMPYDSA-NYEKE-DNMTSHVMDQAINN 364
Qy 310 AISYLGAERPLPYOTPAPTSEMPVPISSVYPTALTRADMP---NGAPOEMEKRRILL 365
Db 365 AINYLGAERPLVQTPPG--SSEVPVVISSMYQHLKPPSGPPRSNHSQAQDANLLLS 423

hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92016-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.
Sequence 431 AA;

Query Match 42.0%; Score 1146; DB 1; Length 431;
Best Local Similarity 51.0%; Pred. NO. 8.1e-95;
Matches 244; Conservative 64; Mismatches 94; Indels 76; Gaps 15;

QY	48	NEDED-----AGEDSMKVKEYSDRDENIMKPEPMGDAEISEMYPYVARIEYSOYESIKLER	103
DB	4	DEGODMSQVSGKEPSPYSDPTDEGDEPMPVEDLSTSGAQ-----	45
QY	104	HVPYDNRSPITSGKMNCDCVCLSCISFNVLVHKRSTGTGPRPFCNOCGASFTQKGNLLRH	163
DB	45	-----QNSKSDRG-----MGERPFCNOCGASFTQKGNLLRH	76
QY	164	IKLHTGKPKFKCHLCNYACQRRDALTGHLURTHSVKPEYKCEFCGRSYKQRSSLEEHEKRC	223
DB	77	IKLHSGEKPKFKCHLCNYACRRRDALTGHLRTHSVGRPHKCGYCGRSYKQRSSLEEHEKRC	136
QY	224	RAFLONPDLGDAASV---EARHLK-----AEMGSEALVLDRLASNVAKRKSSMPQKFTG	275
DB	137	HNLYESMGLPGVCPVKEETNNHMEADLCKIGAERSLVLDRLASNVAKRKSSMPQKFTG	196
QY	276	EKRHCF--DANYNPYMYEKENEMXQTRMMDQANNNAISVLGAERPLVQTPPATPEMVA	334
DB	197	DK--CLSDMPYDSA-NYEKE-DMWTSHVMDQANNNAIVLYGAESRPLVQTPPG--SSEVV	251
QY	335	PVTSISYPIALTRADMP-----MGAOMEKKRILLPEKILPBERGLUSPNNASQDSTDTS	390
DB	252	PVTSISMYQLHKPPSDGPPSRNSHQAQAVONLLLSLAKSVSSERAPSNCSQDSTDTS	311
QY	391	NHEDROH--LYOQSHVYLPQARNGMPLLEKVPESFELLKPPPTCLLRDSITKVNNKEGYMD	448
DB	312	NAEQSRGLIYLNH-INPHARNGL-ALKEEQRAYEVLRAASENSQDARFVYSTSGEQUK	369
QY	449	VFRCDHCHVLFLOYMFTIHM-----GCHGFRDPPECNMGYRSHDRYEFSSHIARGHR	503
DB	370	VYKCEHCRVLFLOHVMYTIHMGCHGCHGFRDPPECNMGYHSDRYEFSSHITRGEHR	427

RESULT 15

R48963	
ID	R46963 standard; Protein; 431 AA.
AC	R46963;
DT	21-OCT-1994 (first entry)
DE	Murine Ikaros zinc finger protein.
DE	Murine Ikaros zinc finger; protein; immune disorder; therapy; treatment;
KW	Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
KW	corpus striatum; regulatory gene.
OS	Mus musculus.
PN	WO9406814-A.
PD	31-MAR-1994.
PF	14-SEP-1993; U08743.
PR	14-SEP-1992; US-946233.
PA	(GEHO) GEN HOSPITAL CORP.
PI	Georgopoulos K;
DR	WPI: 94-118387/14.
DR	N-PSDB; Q44979.
PT	T-cell pathway regulatory gene, Ikaros - encodes family of unique
PT	zinc finger proteins, useful for treating immune system disorders
PS	Claim 14; Page 41-43; 112pp; English.
CC	The Ikaros gene encodes a zinc finger protein which can be used in a
CC	therapeutic composition to treat animals with an immune system
CC	disorder. It may also be used for assessing whether a subject is at
CC	risk for an immune disorder. It is of particular use in treating a
CC	disorder of the corpus striatum.
CC	Sequence 431 AA;
SC	

Search completed: November 6, 1999, 06:55:00
Job time: 82 sec

PS Disclosure; Page 75-76; 102pp; English.
CC The sequence of an Ikaros protein (R92021) is provided in the
CC specification. Ikaros protein is a master regulator of
CC hematopoietic differentiation and a major determinant in lymphocyte
CC differentiation. Isoforms of Ikaros (see R92014-19) arise from
CC differential splicing of Ikaros gene transcripts. Transgenic animals,
CC pref. mice, having a mutated Ikaros transgene, esp. a mutation that
CC alters the DNA binding domain of the Ikaros protein, are used as
CC models to determine the effects of treatments for immune or nervous
CC system disorders.
SQ Sequence 470 AA;

Query Match 44.3%; Score 1208; DB 1; Length 470;
Best Local Similarity 58.3%; Pred. NO. 2.5e-100;
Matches 242; Conservative 50; Mismatches 93; Indels 30; Gaps 11;

QY 111 RPTSGKMNCDVGLGICISFNVLVHKSHTGERPFQCNQCGASFTQKGNLRHKLHTGE 170
DB 60 RLPNGKLKDCIGXCIGPNVLVHKSHTGERPFQCNQCGASFTQKGNLRHKLHSGE 119
QY 171 KPFKCHLCNACRRDALTGHLRTHSVKPKGCGYKQXSLBEKHKRCRAFLQNP 230
DB 120 KPFKCHLCNACRRDALTGHLRTHSVKPKGCGYKQXSLBEKHKRCRCHNLYESM 179
QY 231 DL-GDAASV---EARHTK-----AEMGSEALVLDRLASNVAKKSSMPKFFGEKRCF 281
DB 180 GLPGXXXPVTKETXHXEMAEADCKIGXERSLVLDRLASNVAKKSSMPKFFGLDK-XLS 238
QY 282 DANYNPGYMEKENEMOTRMDDAINAJSYLGAERPLVOTPPAPTSEMPVPISSVY 341
DB 239 DXPYDSA-XYEKEXMMXSHVMDXAINAINLYGAESLRPLVOTPPG-XSEVPVPISPMY 296
QY 342 PIALTRADMPGAPQEMKKR-----ILLPEKILPSERGLSPNNSAQDSTDTSNHE 393
DB 297 QL---HXXSXGPRNSHSAQDAVXXLLLSKAKVXSERERASPSNSCODSTDTSNHE 353
QY 394 DRQH--LYQSHVVLPOARNGMPLLEKVPFSFELLKPPICLRDSIKVINKEGEMDVFR 451
DB 354 EORSGLIYLTNHTX--XXAXXXLKEEXRAYXXLRAASNSQDAXRVVSTSGEQXKVK 411
QY 452 CDHCHVLFLDYVMTIHM---GCHGFDPFECNMGYRSHDRYEFSSHARGHR 503
DB 412 CEHCRLVFLDHVMTIHMXXGCHGFDPFECNMGYHSDRYEFSSHITRGEHR 466

RESULT 12

W72678
ID W72678 standard; Protein; 470 AA.
AC W72678;
DE 14-JAN-1999 (first entry)
KW Ikaros protein general formula.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
KW Alzheimer's disease.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc_difference 1 /note= "any amino acid"
FT Misc_difference 2 /note= "any amino acid"
FT Misc_difference 74 /note= "any amino acid"
FT Misc_difference 163 /note= "any amino acid"
FT Misc_difference 184 /note= "any amino acid"
FT Misc_difference 185 /note= "any amino acid"
FT Misc_difference 186 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"
FT Misc_difference 194 /note= "any amino acid"
FT Misc_difference 196 /note= "any amino acid"
FT Misc_difference 207 /note= "any amino acid"
FT Misc_difference 236 /note= "any amino acid"
FT Misc_difference 240 /note= "any amino acid"
FT Misc_difference 246 /note= "any amino acid"
FT Misc_difference 251 /note= "any amino acid"
FT Misc_difference 252 /note= "any amino acid"
FT Misc_difference 255 /note= "any amino acid"
FT Misc_difference 261 /note= "any amino acid"
FT Misc_difference 285 /note= "any amino acid"
FT Misc_difference 300 /note= "any amino acid"
FT Misc_difference 301 /note= "any amino acid"
FT Misc_difference 302 /note= "any amino acid"
FT Misc_difference 304 /note= "any amino acid"
FT Misc_difference 306 /note= "any amino acid"
FT Misc_difference 316 /note= "any amino acid"
FT Misc_difference 319 /note= "any amino acid"
FT Misc_difference 320 /note= "any amino acid"
FT Misc_difference 329 /note= "any amino acid"
FT Misc_difference 331 /note= "any amino acid"
FT Misc_difference 352 /note= "any amino acid"
FT Misc_difference 367 /note= "any amino acid"
FT Misc_difference 368 /note= "any amino acid"
FT Misc_difference 369 /note= "any amino acid"
FT Misc_difference 371 /note= "any amino acid"
FT Misc_difference 372 /note= "any amino acid"
FT Misc_difference 373 /note= "any amino acid"
FT Misc_difference 374 /note= "any amino acid"
FT Misc_difference 375 /note= "any amino acid"
FT Misc_difference 380 /note= "any amino acid"
FT Misc_difference 384 /note= "any amino acid"
FT Misc_difference 385 /note= "any amino acid"
FT Misc_difference 397 /note= "any amino acid"
FT Misc_difference 407 /note= "any amino acid"
FT Misc_difference 430 /note= "any amino acid"
FT Misc_difference 431 /note= "any amino acid"
FT Misc_difference 432 /note= "any amino acid"

ID R45964 standard; Protein; 537 AA.
AC R45964;
DI 21-OCT-1994 (first entry)
DE Peptide with Ikaros protein activity.
KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
KW corpus striatum; regulatory gene.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 536
FT /note= "Position is encoded by a stop codon in the
FT corresponding nucleotide sequence."
PN WO9406814-A.
PD 31-MAR-1994.
PF 14-SEP-1993; U08743.
PR 14-SEP-1992; US-946233.
PA (GEHO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 94-118387/14.
DR N-PSDB; Q44980.
PT T-cell pathway regulatory gene, Ikaros - encodes family of unique
PT zinc finger proteins, useful for treating immune system disorders
PS Claim 14; Page 44-46; 112pp; English.
CC The Ikaros gene encodes a zinc finger protein which can be used in a
CC therapeutic composition to treat animals with an immune system
CC disorder. It may also be used for assessing whether a subject is at
CC risk for an immune disorder. It is of particular use in treating a
CC disorder of the corpus striatum.
SQ Sequence 537 AA;

Query Match 46.4%; Score 1265.5; DB 1; Length 537;
Best Local Similarity 60.8%; Pred. No. 2.1e-105;
Matches 251; Conservative 60; Mismatches 71; Indels 31; Gaps 12;

QY 111 RPTSGKMDVGLSCISFNVLWVKRSHGTGERPPQCGASFTQKGNLRLHKLHGTGE 170
DB 130 RLPNGKLCDCIGIICIGPNVLWVKRSHGTGERPPQCGASFTQKGNLRLHKLHSGE 189
QY 171 KPFKCHLNCACQRRDALTGHLTHSVGKPKYKCEGCRSYKQSSLEHKKRCRAFLQNP 230
DB 190 KPFKCHLNCACRRRDALTGHLTHSVGKPKYKCEGCRSYKQSSLEHKKRCRCHNLESM 249
QY 231 DL-GDAVSV---EARHK---AEMGSEALVLDRLASNAVKRSMPOKFIGEKRCF 281
DB 250 GLPGTLPVKEETKHSMAEDLCKIGSERSLVLDRLASNAVKRSMPOKFLGDK-GLS 308
QY 282 DANYNPGMYKENEMMOTRMDDQAINNAISYLGAEPRLVQTPAPTSEMVPVSISSVY 341
DB 309 DTPYDSA-TYKENEMKSHVMDQAINNAISYLGAESLRLVQTPPG-GSEVVPVISPMY 366
QY 342 PIALTRADMPGAPQE-----MEKRRILLPEKILPSERGLSPNNNSAQDSTDTDSNHE 393
DB 367 QLH-RRSE---GTPRSNHSADSAVEYLLLSKAKLVPSEREASPSNCSQDSTDTESNNE 422
QY 394 DRQH--LYQQSHVLPQARNMPL-LKEVPSRFELLKPPICLRDSIKVINKEGEVMDVF 450
DB 423 EORSGLTLTNHI-----ARRAQRVSLKEHRAIDYLVRAASNSQDAFRVSTSGEQMKVY 478
QY 451 RCDCHVLFVLDYVMTIIMGCHGFRDPFCNMGCRSHDRYEFFSSHARGEHR 503
DB 479 KCEHCRVLFVLDYVMTIIMGCHGFRDPFCNMGCRSHDRYEFFSSHARGEHR 531

RESULT 11
F92021
ID R92021 standard; Protein; 470 AA.
AC R92021;
DI 09-MAY-1996 (first entry)
DE Ikaros protein.
KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW immunocompromised; immune system disorder; nervous system disorder;
KW animal model.
AC Not specified.

FH Key Location/Qualifiers
FT misc_difference 1..2 /note= "unidentified amino acids"
FT misc_difference 74 /note= "unidentified amino acid"
FT misc_difference 163 /note= "unidentified amino acid"
FT misc_difference 184..186 /note= "unidentified amino acid"
FT misc_difference 194 /note= "unidentified amino acids"
FT misc_difference 196 /note= "unidentified amino acid"
FT misc_difference 207 /note= "unidentified amino acid"
FT misc_difference 236 /note= "unidentified amino acid"
FT misc_difference 240 /note= "unidentified amino acid"
FT misc_difference 246 /note= "unidentified amino acid"
FT misc_difference 251..252 /note= "unidentified amino acids"
FT misc_difference 255 /note= "unidentified amino acid"
FT misc_difference 261 /note= "unidentified amino acid"
FT misc_difference 285 /note= "unidentified amino acid"
FT misc_difference 300..302 /note= "unidentified amino acids"
FT misc_difference 304 /note= "unidentified amino acid"
FT misc_difference 306 /note= "unidentified amino acid"
FT misc_difference 316 /note= "unidentified amino acid"
FT misc_difference 319..320 /note= "unidentified amino acids"
FT misc_difference 329 /note= "unidentified amino acid"
FT misc_difference 331 /note= "unidentified amino acid"
FT misc_difference 352 /note= "unidentified amino acid"
FT misc_difference 367..369 /note= "unidentified amino acids"
FT misc_difference 371..375 /note= "unidentified amino acids"
FT misc_difference 380 /note= "unidentified amino acid"
FT misc_difference 384..385 /note= "unidentified amino acids"
FT misc_difference 397 /note= "unidentified amino acid"
FT misc_difference 407 /note= "unidentified amino acid"
FT misc_difference 430..432 /note= "unidentified amino acids"
FT misc_difference 467 /note= "unidentified amino acid"
FT misc_difference 469 /note= "unidentified amino acid"
FT /note= "unidentified amino acid"
PN WO9604372-A1
PD 15-FEB-1996.
PF 28-JUL-1995; U09345.
PR 28-JUL-1994; US-283300.
PA (GEHO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 96-129389/13.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is
PT severely immuno-compromised and can be used as model to determine
PT effects of treatment for immune and nervous system disorders

CC This is the amino acid sequence of human Ikaros protein isoform
 CC IK-1, deduced from a cDNA clone (see V42806) obtained from a
 CC Jurkat T cell line cDNA library. Native Ikaros is active in the
 CC early stages of lymphocyte differentiation, binding to and
 CC activating the CD3-delta gene enhancer (see V42804). Proteins
 CC of the human Ikaros family (see also W0969 and W0971) are
 CC isoforms that arise from differential splicing of Ikaros gene
 CC transcripts, and contain different combinations of zinc fingers.
 CC They are expressed primarily in T cells in the adult and may play a
 CC role as a genetic switch regulating entry into the T cell lineage.
 CC The human and murine sequences (see also W0963 and W0965-68) are
 CC very similar. The invention provides Ikaros nucleic acids, vectors
 CC and host cells expressing Ikaros polypeptides. These can be used
 CC to treat T and B cell diseases (e.g. immune deficiencies caused by
 CC drugs, radiation or cancer), to control expression of heterologous
 CC genes placed under control of an Ikaros-responsive element, to
 CC treat nervous system diseases (e.g. Alzheimer's disease) and to
 CC modulate cell division, amplification or differentiation, especially
 CC in haematopoietic cells. Some Ikaros isoforms are antagonistic of
 CC others and may be used to inhibit interaction with DNA sequences.
 SQ Sequence 461 AA;

Query Match 46.9%; Score 1279.5; DB 1; Length 461;
 Best Local Similarity 61.5%; Pred. No. 9.4e-107;
 Matches 254; Conservative 59; Mismatches 69; Indels 31; Gaps 12;

QY 111 RPTSGKMNCDVCGLSICISFNVLWVHKRSHTGERPFQCCGASFTQKGNLLRHKLHTE 170
 DB 56 RLPNGKLCDCIGICIGPNVLWVHKRSHTGERPFQCCGASFTQKGNLLRHKLHTE 115
 QY 171 KPFKCHLCNYACQRDALTGHLRTHSVKPYKCFECGRSYKQRSLEHKEKRCRAFLQNP 230
 DB 116 KPFKCHLCNYACRRDALTGHLRTHSVKPYKCFECGRSYKQRTSLEHKEKRCNYLES 175
 QY 221 DL-GDAASV---EARHIK---AEMGSEALVLDRLASNVAKRKSSMPQKFIGEKRCF 281
 DB 176 GLPTGLYPVKEETKSHMAEDLCKIGERSLVLDRLASNVAKRKSSMPQKFLGDK-GLA 234
 QY 282 DANYNPGYMEKENEMMOTRMMDQAINNAISYLGAFAFRPLVQTPTPTSEMVPVSISSY 341
 DB 235 DTPYDSA-TYEKENEMKSHVMDQAINNAISYLGAFAFRPLVQTPTPTG-GSEVPVISP 292
 QY 342 PIALTRADMPGARQE-----WEKRIILLPEKILPSERGLSPNNSAQDSTDTDSNHE 393
 DB 293 QLH-RRSE---GTPRSNHSQDSAVEYLLLSKALVPSEREASPSNSCQDSTDTESNNE 348
 QY 394 DRQH--LYQQSHVLPQARNMPL-LKEVPRSFELLKPPICLRDSIKVINKEGEVMDVF 450
 DB 349 EQRSLIYLTNHI---ARRAQRVSLKEEHRAYDILLRAENSQDALRVVSTSGEQMKVY 404
 QY 451 RCDHCHVLFVDYVMTIHMCGHGRDPPECNMGYSRSHDRYEFSSHIARGEHR 503
 DB 405 KCEHCRVLFVDHVMYTIHMCGHGRDPPECNMGYSRSHDRYEFSSHIARGEHR 457

RESULT 7
 R46965 ID R46965 standard; Protein; 568 AA.
 AC R46965;
 DT 21-OCT-1994 (first entry)
 DE Ikaros zinc finger protein isoform IK-1.
 KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
 KW corpus striatum; regulatory gene.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT region 1..53
 FT /label= Exons 1/2.
 FT region 54..141
 FT /label= Exon 3.
 FT region 142..247
 FT /label= Exon 4.
 FT region 248..288

FT region /label= Exon 5.
 FT 289..333
 FT /label= Exon 6.
 FT 334..568
 FT /label= Exon 7.
 PN W09406814-A.
 PD 31-MAR-1994.
 PF 14-SEP-1993; U08743.
 PR 14-SEP-1992; US-946233.
 PA (GENO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 94-118387/14.
 PT T-cell pathway regulatory gene, Ikaros - encodes family of unique
 PT zinc finger proteins, useful for treating immune system disorders
 PS Claim 14; Figure 4; 112pp; English.
 CC The Ikaros gene encodes a zinc finger protein which can be used in a
 CC therapeutic composition to treat animals with an immune system
 CC disorder. It may also be used for assessing whether a subject is at
 CC risk for an immune disorder. It is of particular use in treating a
 CC disorder of the corpus striatum.
 SQ Sequence 568 AA;

Query Match 46.9%; Score 1279; DB 1; Length 568;
 Best Local Similarity 49.4%; Pred. No. 1.4e-106;
 Matches 278; Conservative 72; Mismatches 131; Indels 82; Gaps 18;

QY 12 STERQPLTPSPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVKDEYSDRDENI 71
 DB 13 SGRESPPVSDTPDE-GDEPMPVPEDLSTTSQ--AQSKSDRGNAS-NVKVETQSDENG 68
 QY 72 MKPEPMG-----DAESEMPSYAREYSDYES----- 99
 DB 69 RACEMNGEECAEDLRLMLDASGEKMGSHRDQGSALSQGVGIRLPNGKLCDCIGIVCIG 128
 QY 99 --IKLHRHVPYDNRSP-----TSCKMNCDCVGLSCISFNVLWVHKRS 138
 DB 129 PNVLWVHKRSHTGERPFQCCGASFTQKGNLLRHKLHTEKRCRAFLQNP 188
 QY 139 HTGERPFQCCGASFTQKGNLLRHKLHTEKRCRAFLQNP 198
 DB 189 HTGERPFQCCGASFTQKGNLLRHKLHTEKRCRAFLQNP 248
 QY 199 KPYKCFECGRSYKQRSLEHKEKRCRAFLQNP---PDLGDAASVEARHIK-----AEMGSE 250
 DB 249 KPHKCGYCGRSYKQRSLEHKEKRCRAFLQNP---PDLGDAASVEARHIK-----AEMGSE 308
 QY 251 RALVLDRLASNVAKRKSSMPQKFIGEKRCF--DANYNPGYMEKENEMMOTRMMDQAINN 309
 DB 309 RSLVLDRLASNVAKRKSSMPQKFIGEKRCF--DANYNPGYMEKENEMMOTRMMDQAINN 364
 QY 310 AISYLGAFAFRPLVQTPTPTSEMVPVSISSYPIALTRADMP-----MGAPQMEKKRILL 365
 DB 365 AINYLGAESLRPLVQTPTPTG-GSEVPVSISSYPIALTRADMP-----MGAPQMEKKRILL 423
 QY 366 PEKILPSERGLSPNNSAQDSTDTDSNHEDRQH--LYQQSHVLPQARNMPLKEVPRSF 423
 DB 424 KAKSVSSEREASPSNSCQDSTDTESNAEQRSGLIYLTNH-INPHARNGL-ALKEEQRAY 481
 QY 424 ELLKPPICLRDSIKVINKEGEVMDVFCDCRCHVLFVDYVMTIHM--GCHGRDPDFEC 480
 DB 482 EVLRAAENSQDALRVVSTSGEQLKVKYKCHRCRVLFVDHVMYTIHMCGHGRDPDFEC 541
 QY 481 NMCGYSRSHDRYEFSSHIARGEHR 503
 DB 542 NMCGYSRSHDRYEFSSHIARGEHR 564

RESULT 8
 R92015 ID R92015 standard; Protein; 461 AA.
 AC R92015;
 DT 09-MAY-1996 (first entry)

QY 407 POARNGMLLKEVPRSEFLLKPPICLRDSIKVINKEGEVMDVFCRDCHVFLDYVMT 466
 DB 416 PHARNGI-ALKEQRAYEVLRASSENSQDAFRVSTSGEQLKYKCHCRVFLDHYMT 474
 QY 467 IHM---GCHGFDPFCNMGCGYRSHDRYFESSHARGHR 503
 DB 475 IHMGCHGCHGFDPFCNMGCGYHSQDRYFESSHITRGEHR 514

RESULT 5

W70966 ID W70966 standard; Protein; 518 AA.
 AC W70966; 03-MAR-1998.
 DT 11-JAN-1999 (first entry)
 DE Mouse Ikaros isoform m1k-1.
 KW Ikaros; m1k-1; transcription factor; mouse; lymphocyte;
 cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT Region 119..139
 FT /note= "zinc finger motif"
 FT Region 147..167
 FT /note= "zinc finger motif"
 FT Region 175..195
 FT /note= "zinc finger motif"
 FT Region 203..224
 FT /note= "zinc finger motif"
 FT Region 460..480
 FT /note= "zinc finger motif"
 FT Region 491..513
 FT /note= "zinc finger motif"
 FT CA2194256-A.
 PN 03-MAR-1998.
 PD 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR N-PSDB; V42808.
 PT New nucleic acid encoding Ikaros protein involved in early
 differentiation of lymphocytes - existing in several isoforms, and
 related products, used to treat e.g. immune diseases or cancer and
 to control cell differentiation
 PS Claim 7; Page 75-77; 158pp; English.
 CC This is the amino acid sequence of murine Ikaros protein isoform
 m1k-1, deduced from a cDNA clone (see V42808) obtained from a
 mature murine T cell line E14 library. Native Ikaros is active
 in the early stages of lymphocyte differentiation, binding to and
 activating the CD3-delta gene enhancer (see V42804). Proteins
 of the murine Ikaros family (see also W70963 and W70965-68) are
 isoforms that arise from differential splicing of Ikaros gene
 transcripts, and contain different combinations of zinc fingers.
 They are expressed primarily in T cells in the adult and may play a
 role as a genetic switch regulating entry into the T cell lineage.
 CC The murine and human sequences (see W70964, W70969 and W70971) are
 very similar. The invention provides Ikaros nucleic acids, vectors
 and host cells expressing Ikaros polypeptides. These can be used
 to treat T and B cell diseases (e.g. immune deficiencies caused by
 CC drugs, radiation or cancer), to control expression of heterologous
 genes placed under control of an Ikaros-responsive element, to
 CC treat nervous system diseases (e.g. Alzheimer's disease) and to
 CC modulate cell division, amplification or differentiation, especially
 CC in haematopoietic cells. Some Ikaros isoforms are antagonistic of
 CC others and may be used to inhibit interaction with DNA sequences.
 SQ Sequence 518 AA;

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEEQPLPTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGEDSMKVKDEYSRDEMI 71
 DB 13 SGKESPPVSDTPE-GDEMPVPEDLSTTSG--AQONSXSDRGMAS-NYKVTQSDSENG 68
 QY 72 MKPEPMG-----DAESEMPTYSYAREYSYIESIKLERHVPYDNSRPTSGKNKCDV 121
 DB 69 RACEMNGEBCAEDRLMDASGEKMGSHRDQSSALS-----GVGGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHRSHTGERPFCNQCQASFTQKGNLLRHIKLTGKPKFKCHLCNYA 181
 DB 122 CGIVCIGPNVLMVHRSHTGERPFCNQCQASFTQKGNLLRHIKLTGKPKFKCHLCNYA 181
 QY 182 CORRDALTGLRTHSVKPKYKCFGCRSYKQSSLEEKERCAFQNLQDGAASV--- 239
 DB 182 CRRRDALTGLRTHSVGPKHKGCGRSYKQSSLEEKERCHNYLESMLGPGVCPVIKE 241
 QY 239 EARHIK-----AEMGSEALVLDRLASNVAKRKSMPQKFIGEKHCF-DANYNPGMYE 292
 DB 242 ETNHNEMAEEDLCKIGAERSLVLDRLASNVAKRKSMPQKFLGDK--CLSDMPYDSA-NYE 298
 QY 293 KENEMQTRMDOAINNAISYLGAERPLVQTPAPTSEMPVVISSVPIALTRADMP- 352
 DB 299 KE-DMTSHVMDQAINNAISYLGAESRLPLVQTPPG-SSEVPVVISSMYQLHKPPSDGPP 356
 QY 352 ---MGAPOQMEKKRILLPEKILPSERGLSPNNSAODSTDTSNHHEDRQH--LYQOSHVYL 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSERASPSQSDSTESNAEEQRSLIYLTNH--IN 415
 QY 407 POARNGMLLKEVPRSEFLLKPPICLRDSIKVINKEGEVMDVFCRDCHVFLDYVMT 466
 DB 416 PHARNGI-ALKEQRAYEVLRASSENSQDAFRVSTSGEQLKYKCHCRVFLDHYMT 474
 QY 467 IHM---GCHGFDPFCNMGCGYRSHDRYFESSHARGHR 503
 DB 475 IHMGCHGCHGFDPFCNMGCGYHSQDRYFESSHITRGEHR 514

RESULT 6

W70964 ID W70964 standard; Protein; 461 AA.
 AC W70964;
 DT 11-JAN-1999 (first entry)
 DE Human Ikaros isoform h1k-1.
 KW Ikaros; h1k-1; transcription factor; human; lymphocyte;
 cell differentiation; T cell; cancer; immunodeficiency;
 KW Alzheimer's disease; therapy; diagnosis.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 64..84
 FT /note= "zinc finger motif"
 FT Region 92..112
 FT /note= "zinc finger motif"
 FT Region 120..140
 FT /note= "zinc finger motif"
 FT Region 148..169
 FT /note= "zinc finger motif"
 FT Region 406..426
 FT /note= "zinc finger motif"
 FT Region 434..456
 FT /note= "zinc finger motif"
 FT CA2194256-A.
 PN 05-MAR-1998.
 PD 02-JAN-1997; 194256.
 PR 05-SEP-1996; US-711417.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-378292/33.
 DR N-PSDB; V42806.
 PT New nucleic acid encoding Ikaros protein involved in early
 differentiation of lymphocytes - existing in several isoforms, and
 related products, used to treat e.g. immune diseases or cancer and
 to control cell differentiation
 PS Claim 7; Page 70-72; 158pp; English.

PN W09604372-A1.
 PD 15-FEB-1996.
 PF 28-JUL-1995; U09345.
 PR 29-JUL-1994; US-283300.
 PA (GEOH) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 96-129389/13.
 DR N-PSDB; T016062.
 PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
 PT Disclosure: Fig 4; 102pp; English.
 PS The sequence of 57.5 kba mouse Ikaros protein mik-1 (R92017) was deduced from mouse Ikaros cDNA (T16062) isolated from a mature T-cell line E15 library. Ikaros protein is a master regulator of hematopoietic differentiation and a major determinant in lymphocyte differentiation. Other isoforms of Ikaros (see R92014, R92016 and R92018-19) arise from differential splicing of Ikaros gene transcripts. Transgenic animals, esp. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.
 CC Sequence 518 AA;
 SQ

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEOPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVDEYSDRDENI 71
 DB 13 SKGESPPVSDTPDE-GDEPMPVPEDLSTSG--AQNSKSDRGMAS-NVKVETOSDEENG 68
 QY 72 MKPEPMG-----DAESEMPSYAREYSDYESIKLERHVPYDINSRPTSGKMNCDV 121
 DB 69 RACENMGEECAEDLRMLDASGEKMGSHRQDGSALS-----GVGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGKPFKCHLCNYA 181
 DB 122 CGVICGPNVLVHKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGKPFKCHLCNYA 181
 QY 182 CORRDLATGHLRTHSVKPYKCEFCGRSYKORSLSLEHKEKRCRAFLQNPDLGDAASV--- 239
 DB 182 CRRDLATGHLRTHSVKPYKCEFCGRSYKORSLSLEHKEKRCRAFLQNPDLGDAASV--- 239
 QY 239 EARHIK-----AEMGSEALVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYME 292
 DB 242 ETNNHMAEDLCKIGAERSLVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYME 292
 QY 293 KENEMQTRMDQAINNAISYLGAERPLVQTPPTAPPTSEMVPVSSVYPIALTRADMP- 352
 DB 299 KE-DNMTSHVMDQAINNAISYLGAERPLVQTPPTAPPTSEMVPVSSVYPIALTRADMP- 352
 QY 352 ---MGAPOEMEKKRILLPEKILPSERGLSPNNSAODSTDSDSNHEDROH--LYQOSHVL 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSQDSTDSDSNHEDROH--LYQOSHVL 406
 QY 407 POARNMPLLAKEVRSFELLPPPTCLDRSDIKVINKEGVMDVDFCDCHVLFIDYVNF 466
 DB 416 PHARNGL-ALKEEQRAYEVLRAAENSODAFRVVSTSGEQLKVKYKCEHCRVFLDHVMT 474
 QY 467 IHM---GCHGRDPPECNMGYRSHDRYEFSSHIARGHR 503
 DB 475 IHMGCHGCHGRDPPECNMGYRSHDRYEFSSHIARGHR 514

RESULT 4
 W72674
 ID W72674 standard; Protein; 518 AA.
 AC W72674;
 DT 14-JAN-1999 (first entry)
 DE Mouse Ikaros mik-1.
 KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;

KW differentiation marker; immune system; corpus striatum; AIDS;
 KW Alzheimer's disease.
 OS Mus sp.
 PN US5824770-A.
 PD 20-OCT-1998.
 PF 05-JUN-1995; 465590.
 PR 02-MAY-1994; US-238212.
 PR 14-SEP-1992; US-946233.
 PR 14-SEP-1993; US-121438.
 PR 05-JUN-1995; US-465590.
 PA (GEOH) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 98-582821/49.
 DR N-PSDB; V66971.
 PT Ikaros poly-peptide(s) - useful for treating disorders of immune system or corpus striatum
 PS Claim 1: Column 61-66; 111pp; English.
 CC The present invention describes a purified peptide having at least one of the following properties: (a) it stimulates transcription of a DNA sequence under the control of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (c) it competitively inhibits the binding of a naturally occurring Ikaros isoform to any of a delta A element, an NFkB element or an Ikaros binding oligonucleotide consensus sequence; (d) it competitively inhibits Ikaros binding to Ikaros responsive elements; or (e) it inhibits protein-protein interactions of transcriptional complexes formed with naturally occurring Ikaros isoforms. The proteins, provided that they stimulate gene transcription under the control of delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit binding of naturally occurring Ikaros isoforms to delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides, competitively inhibit Ikaros binding to Ikaros-responsive elements and/or inhibit protein-protein interactions of transcriptional complexes with naturally occurring Ikaros isoforms, can be used to treat immune system disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g. Alzheimer's disease. The present sequence represents a specifically claimed mouse Ikaros protein.
 CC Sequence 518 AA;
 SQ

Query Match 47.1%; Score 1284; DB 1; Length 518;
 Best Local Similarity 53.3%; Pred. No. 4.4e-107;
 Matches 277; Conservative 68; Mismatches 129; Indels 46; Gaps 17;

QY 12 STEOPLTESPDALNDYSLPKPHEIENVDSREAPANEDEDAGDSMKVDEYSDRDENI 71
 DB 13 SKGESPPVSDTPDE-GDEPMPVPEDLSTSG--AQNSKSDRGMAS-NVKVETOSDEENG 68
 QY 72 MKPEPMG-----DAESEMPSYAREYSDYESIKLERHVPYDINSRPTSGKMNCDV 121
 DB 69 RACENMGEECAEDLRMLDASGEKMGSHRQDGSALS-----GVGIRLPNGKLKCDI 121
 QY 122 CGLSCISFNVLVHKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGKPFKCHLCNYA 181
 DB 122 CGVICGPNVLVHKRSHGTGERPFCNOCQSGASFTQGNLLRHKLHGTGKPFKCHLCNYA 181
 QY 182 CORRDLATGHLRTHSVKPYKCEFCGRSYKORSLSLEHKEKRCRAFLQNPDLGDAASV--- 239
 DB 182 CRRDLATGHLRTHSVKPYKCEFCGRSYKORSLSLEHKEKRCRAFLQNPDLGDAASV--- 239
 QY 239 EARHIK-----AEMGSEALVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYME 292
 DB 242 ETNNHMAEDLCKIGAERSLVLDRLASNAVKRSMQPKFGEKRRHCF-DANYNPGYME 292
 QY 293 KENEMQTRMDQAINNAISYLGAERPLVQTPPTAPPTSEMVPVSSVYPIALTRADMP- 352
 DB 299 KE-DNMTSHVMDQAINNAISYLGAERPLVQTPPTAPPTSEMVPVSSVYPIALTRADMP- 352
 QY 352 ---MGAPOEMEKKRILLPEKILPSERGLSPNNSAODSTDSDSNHEDROH--LYQOSHVL 406
 DB 357 RSNHSAQDAVDNLLLSKAKSVSSERASPSNSQDSTDSDSNHEDROH--LYQOSHVL 406
 QY 407 POARNMPLLAKEVRSFELLPPPTCLDRSDIKVINKEGVMDVDFCDCHVLFIDYVNF 466
 DB 416 PHARNGL-ALKEEQRAYEVLRAAENSODAFRVVSTSGEQLKVKYKCEHCRVFLDHVMT 474
 QY 467 IHM---GCHGRDPPECNMGYRSHDRYEFSSHIARGHR 503
 DB 475 IHMGCHGCHGRDPPECNMGYRSHDRYEFSSHIARGHR 514


```
QY 301 RMDQAINNAISYLGAEAFRLVOTPPAPTSEMPVPISSVYPIALTRADMPGAPQEMEK 350
Db 301 RMDQAINNAISYLGAEAFRLVOTPPAPTSEMPVPISSVYPIALTRADMPGAPQEMEK 360
QY 361 KRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRQHLVQOQSHVVLPOARNGMPLLEKVP 420
Db 361 KRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRQHLVQOQSHVVLPOARNGMPLLEKVP 420
QY 421 RSEFLLKPPICLRDSIKVINKGEVMDVPRCDHCHVFLFDYVNFTHMCHGFRDPFEC 480
Db 421 RSEFLLKPPICLRDSIKVINKGEVMDVPRCDHCHVFLFDYVNFTHMCHGFRDPFEC 480
QY 481 NMGCYRSHDRYEFSSHTARGEHRAMLK 507
Db 481 NMGCYRSHDRYEFSSHTARGEHRAMLK 507

RESULT 2
ID W70971 standard; Protein; 516 AA.
AC W70971;
DT 11-JAN-1999 (first entry)
DE Human Ikaros isoform h1k-1.
KW Ikaros; h1k-1; transcription factor; human; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis.
OS Homo sapiens.
FH Key
FT Region
FT 119..139
FT /note= "zinc finger motif"
FT 147..167
FT /note= "zinc finger motif"
FT 175..195
FT /note= "zinc finger motif"
FT 203..224
FT /note= "zinc finger motif"
FT 461..481
FT /note= "zinc finger motif"
FT 489..511
FT /note= "zinc finger motif"
CA2194256-A.
PD 05-MAR-1998.
PR 02-JAN-1997; 194256.
PR 05-SEP-1996; US-711417.
PA (GEO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR N-PSDB; V42840.
PT New nucleic acid encoding Ikaros protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer and
PT to control cell differentiation
PS Claim 1; Page 127-129; 158pp; English.
CC This is the amino acid sequence of human Ikaros protein isoform
CC h1k-1, deduced from a cDNA clone (see V42840) obtained from a
CC Jurkat T cell line cDNA library. Native Ikaros is active in the
CC early stages of lymphocyte differentiation, binding to and
CC activating the CD3-delta gene enhancer (see V42804). Proteins
CC of the human Ikaros family (see also W70964 and W70969) are
CC isoforms that arise from differential splicing of Ikaros gene
CC transcripts, and contain different combinations of zinc fingers.
CC They are expressed primarily in T cells in the adult and may play a
CC role as a genetic switch regulating entry into the T cell lineage.
CC The human and murine sequences (see also W70963 and W70965-68) are
CC very similar. The invention provides Ikaros nucleic acids, vectors
CC and host cells expressing Ikaros polypeptides. These can be used
CC to treat T and B cell diseases (e.g. immune deficiencies caused by
CC drugs, radiation or cancer) to control expression of heterologous
CC genes placed under control of an Ikaros-responsive element, to
CC treat nervous system diseases (e.g. Alzheimer's disease) and to
CC modulate cell division, amplification or differentiation, especially
CC in hematopoietic cells. Some Ikaros isoforms are antagonistic of
CC others and may be used to inhibit interaction with DNA sequences.
```

```
SQ Sequence 516 AA;
Query Match 47.7%; Score 1302; DB 1; Length 516;
Best Local Similarity 52.3%; Pred. No. 1.1e-108;
Matches 275; Conservative 75; Mismatches 124; Indels 52; Gaps 16;
QY 8 VELKSTEEQPLTESPDALNDYSLPKPHEIENVDSREAPANEDEAGEDSMKVKDYSDR 67
Db 9 MSFSSGKESPPVSDTPDE-GDEPMPITPEDLTSTSGQSSKSDRVVASN---VKVETQSD 64
QY 68 DENIMKPEPMG-----DAESEMPSYAREYSDYESIKLERHHPYDINSRPSGKM 117
Db 65 EENGACEMNGEACEDLRLMDASGEKMGSHRDGSSALS-----GVGGIRLPNGKL 117
QY 118 NCDVCGLSGISFNVLVHVKRSHTGERPFCNQCNGASFTQGNLLRHLRIKLTGEKPFKCHL 177
Db 118 KCDICGIIICIGPNVLVHVKRSHTGERPFCNQCNGASFTQGNLLRHLRIKLSGKPFKCHL 177
QY 178 CNYACORRDALTGHLRTHSVKPYKCEFCGRSYKQSSLEEHKERCRAFLQNPDL-GDAA 236
Db 178 CNYACRRRDALTGHLRTHSVGPKHKGCGYKGRSYKORTSLEEHKERCNHYLESGLPGTLY 237
QY 237 SV---EARHIK-----AEMGSRALVLDRLASNVAKRKSSMPQKFIGEKRFCDANYNG 288
Db 238 PVIKEETHKSEMAEDLCKIGSERSVLVLDRLASNVAKRKSSMPQKFLGDK-GLSDTTPDSA 296
QY 289 YMYKENENMOTRMMDQAINNAISYLGAEAFRLVOTPPAPTSEMPVPISSVYPIALTRA 348
Db 297 -TYEKENENMKSHVMDQAINNAISYLGAEASRLVOTPPG-GSEVVPVISPVQLH-RSS 353
QY 349 DMPMGAPQE-----MEKKRILLPEKILPSERGLSPNNSAODSTDSDNHNEDRQHL-L 398
Db 354 E--GTPSRNHSQAQDSAVEYLLLLSKALVPSEREASPSNSCQSDSTDTSNNEEQSGLI 410
QY 399 YQOSHVVLPQARNGMPL-LKEVPRSFELKPPICLRDSIKVINKGEVMDVPRCDHCHV 457
Db 411 YLTNHI-----ARRAQRVSLKEEHRAVDLLRAASENSQDALRVVSTSGEOMKVKCEHCRV 466
QY 458 LFLDYVMFTIHMCHGFRDPFECNMGYRSHDRYEFSSHTARGEHR 503
Db 467 LFLDHVMYTIHMCHGFRDPFECNMGYRSHDRYEFSSHTARGEHR 512

RESULT 3
R92017 standard; Protein; 518 AA.
AC R92017;
DT 09-MAY-1996 (first entry)
DE Murine Ikaros protein m1k-1.
KW Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW immunocomprised; immune system disorder; nervous system disorder;
KW animal model; m1k-1.
OS Mus musculus.
FH Key
FT domain
FT 119..140
FT /label= F1
FT /note= "zinc finger domain F1"
FT 147..167
FT /label= F2
FT /note= "zinc finger domain F2"
FT 175..195
FT /label= F3
FT /note= "zinc finger domain F3"
FT 203..224
FT /label= F4
FT /note= "zinc finger domain F4"
FT 460..480
FT /label= F5
FT /note= "zinc finger domain F5"
FT 491..513
FT /label= F6
FT /note= "zinc finger domain F6"
```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 06:53:38 ; Search time 21.23 Seconds
(without alignments)
565.655 Million cell updates/sec

Title: US-09-019-348-2
Perfect score: 2728
Sequence: 1 MEDIQPTVELKSTEEQPLPT.....HDIYFSSHTARGEHRMLK 507

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_35:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2724	99.9	507	W15574	Mouse Aiolos polyp
2	1302	47.7	516	W70971	Human Ikaros isofo
3	1284	47.1	518	R92017	Murine Ikaros prot
4	1284	47.1	518	W72674	Mouse Ikaros mik-1
5	1284	47.1	518	W70966	Mouse Ikaros isofo
6	1279.5	46.9	461	W70964	Human Ikaros isofo
7	1279	46.9	568	R46965	Ikaros zinc finger
8	1276.5	46.8	461	R92015	Human Ikaros prote
9	1276.5	46.8	461	W72672	Human Ikaros, Ikar
10	1265.5	46.4	537	R46964	Peptide with Ikar
11	1208	44.3	470	R92021	Ikaros protein, Tr
12	1208	44.3	470	W72678	Ikaros protein gen
13	1205	44.2	470	W70970	Ikaros isoform 1 c
14	1146	42.0	431	R92014	Murine Ikaros prot
15	1143	41.9	431	R46963	Murine Ikaros zinc
16	1143	41.9	431	W72671	Murine Ikaros mik-
17	1143	41.9	431	W70963	Mouse Ikaros isofo
18	1024.5	37.6	209	W15575	Human Aiolos polyp
19	1020.5	37.4	334	R92020	Ikaros protein, Tr
20	1020.5	37.4	334	W72677	Ikaros isoform. Ik
21	1020.5	37.4	334	W70969	Human Ikaros prote
22	999.5	36.6	390	R92018	Murine Ikaros prot
23	999.5	36.6	390	W72675	Mouse Ikaros mik-4
24	997.5	36.6	390	W70967	Mouse Ikaros isofo
25	993	36.4	432	R92016	Murine Ikaros prot
26	993	36.4	432	W72673	Mouse Ikaros mik-3
27	993	36.4	432	W70965	Mouse Ikaros isofo
28	675	24.7	376	W70968	Murine Ikaros prot
29	671	24.6	376	R92019	Mouse Ikaros isofo
30	671	24.6	376	W72676	Mouse Ikaros mik-5
31	600	22.0	236	W72680	Ikaros protein SEQ
32	588.5	21.6	236	W72679	Ikaros protein SEQ
33	537.5	19.7	238	W72681	Ikaros protein SEQ
34	313	11.5	56	W72688	Ikaros protein SEQ
35	313	11.5	56	W72689	Ikaros protein SEQ
36	303	11.1	56	W72687	Ikaros protein SEQ
37	273.5	10.0	532	W52187	Human BCRAL-associ
38	273	10.0	711	W37504	Human OTK18. Novel
39	270.5	9.9	561	R82404	Human SRE-2BP anal
40	269	9.9	412	Y07068	Renal cancer assoc
41	265.5	9.7	706	R68743	BCL-6 zinc finger
42	261.5	9.6	803	W81756	Myc-binding zinc-f
43	256.5	9.4	181	R83490	Zif268-Zif268 zinc

44 256.5 9.4 181 1 W84394 The zif268-zif268
45 253 9.3 688 1 W06316 TRP-1 protein. DNA

ALIGNMENTS

RESULT 1
W15574 ID W15574 standard; Protein; 507 AA.
AC W15574;
DT 07-JUL-1997 (first entry)
DE Mouse Aiolos polypeptide.
KW Aiolos; transcription activator; immune system; T lymphocyte;
KW B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;
KW transgenic animal.
OS Mus sp.
FH Key Location/Qualifiers
FT domain 290..344
FT /label= Activation_domain
FT /note= "conserved activation domain of Aiolos"
PN W09714714-AT.
DP 24-APR-1997.
PF 17-OCT-1996; U16774
PR 18-OCT-1995; US-005529.
PR 14-MAY-1996; US-017646.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K, Morgan BA;
DR WPI: 97-245047/22.
DR N-PSDB; T60490
PT Aiolos polypeptide and corresponding DNA - used to reconstitute a
PT mammalian immune system, for the treatment of T cell leukaemia(s)
PT and lymphoma(s)
PS Disclosure: Page 76-78; 115pp; English.
CC Mouse Aiolos (W15574) is a homologue of Ikaros whose expression is
CC restricted to lymphoid lineage. Its amino acid sequence was
CC deduced from a cDNA clone (T60490) obtd. from a mouse spleen cDNA
CC library. A human Aiolos partial polypeptide has also been
CC identified (see also W15575). Aiolos forms homodimers that are
CC potent transcriptional activators. It also forms less potent
CC heterodimers with Ikaros. Recombinant Aiolos can be produced in
CC transformed host cells. Aiolos misexpressing immune system
CC components can be used to reconstitute the immune system of a
CC mammal. The Aiolos protein, coding sequence or cells expressing
CC Aiolos may be used to treat a disorder in an animal, esp. by gene
CC therapy. Such disorders include T-cell leukaemias and lymphomas.
SQ Sequence 507 AA;

Query Match 99.9%; Score 2724; DB 1; Length 507;
Best Local Similarity 99.8%; Pred. No. 2.5e-236;
Matches 506; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEDIQPTVELKSTEEQPLPTESPDALNDYSLPKPHEINVDNSREAPANEDEDAGEDSMKV 60
DB 1 MEDIQPTVELKSTEEQPLPTESPDALNDYSLPKPHEINVDNSREAPANEDEDAGEDSMKV 60
QY 61 KDEYSDRDENIMKPEPMGDAEEMPSYAREYSDYESIKLERHVPYDNRPTSGKNMCD 120
DB 61 KDEYSDRDENIMKPEPMGDAEEMPSYAREYSDYESIKLERHVPYDNRPTSGKNMCD 120
QY 121 VCGLSCISFNVLVHKRSHTGERPFOCNOCCGASFTQKGNLLRHKLHTGKPKFKCHLCNY 180
DB 121 VCGLSCISFNVLVHKRSHTGERPFOCNOCCGASFTQKGNLLRHKLHTGKPKFKCHLCNY 180
QY 181 ACQRDALTGHLRTHSVKEPKYKCEFCGRSRYKQSSLSSEHKERCRAFTLQNDLGDAAASVEA 240
DB 181 ACQRDALTGHLRTHSVKEPKYKCEFCGRDYSKQSSLSSEHKERCRAFTLQNDLGDAAASVEA 240
QY 241 RHKAEMGSRALVLDRLASNAVKRKSMPQKFTIGEKRHCFDANYPGYMYEKENEMMOT 300
DB 241 RHKAEMGSRALVLDRLASNAVKRKSMPQKFTIGEKRHCFDANYPGYMYEKENEMMOT 300

THIS PAGE BLANK (uspto)

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15,
TGTTACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 81 a 96 c 77 g 102 t
ORIGIN

Query Match 4.2%; Score 84; DB 38; Length 356;
Best Local Similarity 73.0%; Pred. No. 1.8e-10;
Matches 108; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 932 gcgctcacgggacacaccttaggacacattctgtggagaagccgtacaaagtgtgagttctgc 991
|| ||||| ||||| ||||| ||||| ||||| || || ||||| || |||||
Db 1 GCTCTCACAGGACACTCAGGACCCATTCTGTGGTAAACCTCACAAAGTGAACACTACTGT 60
QY 992 ggaagaagctacaagcagagaagctccctggaggagcacaaggaacgctgccgagctttt 1051
|| ||||| ||||| || || ||||| || ||||| ||||| ||||| |||||
Db 61 GCCCGAAGCTACAAGCAGCGCAGCTCACTGGAGGAACACAAAGGACCGCTGTCACAACTAT 120
QY 1052 cttcagaaccctgacctgggggacgctg 1079
|| ||||| || ||||| || |||||
Db 121 CTCCACAATGTCAGCATGGAGGCTGCCG 148

Search completed: November 6, 1999, 08:16:48
Job time: 501 sec

University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 129 a 67 c 88 g 69 t 2 others

ORIGIN

Query Match 4.3%; Score 85.2; DB 24; Length 355;
Best Local Similarity 60.3%; Pred.No.9.5e-11;
Matches 141; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 764 gtcttgatggttcataagcgaaaccacataccgcggcgaacgccgttcccagtgtaatcagtcg 823
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 36 GGCTGAGGAACATCATAGAACAACACACACAGGGGAGAACCCCTACAAATGTATCACTGT 95
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 824 ggggatctttactcagaagaagtaacctctcgcgtcatattaaaactgcacacaggggaaa 883
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 96 GGAAGAGCTTCGGTGCAGAAATCACAACTCAGAGGACATCATAGAATTACACAGGGGAA 155
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 884 aaacccctttaagtgtcacctcttcaactacgcacgacaaaggagagatgcgtcacgga 943
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 156 AAACCCTATAAATGTAATCATTTGTGGGAAGCITTTCAGTCAGAAATCAAACCTCAGAGTA 215
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 944 caocttaggacacattctgtggagagccgcacaaagtgtgagttctcggaaga 997
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 CATCAGAAACTCATATCTGGGGAGAAAACCCCTATCAATGTGAGGAGTGCGAAAA 269
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15

AA764246

LOCUS v449e05.r1 Soares 2NDMT Mus musculus cdna clone IMAGE:1225760 5'
similar to TR:O13100 O13100 IKAROS HOMOLOG ; mRNA sequence.

ACCESSION AA764246

NID q2811768

VERSION AA764246.1 GI:2811768

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubouque T.,
Geisler S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
Schellenberg K., Stepien M., Tan F., Underwood K., Moore B.,
Theisinger B., Wylie T., Lennon G., Soares B., Wilton R. and
Waterston R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

TITLE On May 8, 1995 this sequence version replaced gi:801088.

JOURNAL Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:651352

COMMENT Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 340.
Location/Qualifiers
1. .356
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1225760"
/clone_lib="Soares 2NDMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

FEATURES source


```

BASE COUNT      Soares and M.Fatima Bonaldo."
ORIGIN          124 a   162 c   152 g   171 t

Query Match      10.3%; Score 204; DB 40; Length 609;
Best Local Similarity 93.1%; Pred. No. 2.9e-39;
Matches 257; Conservative 0; Mismatches 15; Indels 4; Gaps 4;

Qy 690 cctatgacaaag-cagacacacagtggaagatgaactgcagctgtgcggttatcc 748
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 CCTTTGACAAACAGCCAGACACACCGTGGAAAGATGAAGTGGAGCTGTGGGTTATCCC 550

Qy 749 tgcattagcttc-aacgtcttgatggttcac-aagcgaagccatcccggaaccccg 806
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 549 TGCATTAGCTTCAAAAGCTGTTGATGGTTCTATAAAGCGAAGCCATACCCGCGCCCGTT 490

Qy 807 tccagttaactcagtcgaggggcatctttactcagaaggaagtaacctcctcgatatta 866
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 CCACTGTATCAGTGC-GGGCATCTTTTACTCAGAAAGGTAACTCTCCGTCATATTA 431

Qy 867 aactgcacacgggggaaacacctttaagtgtcacctctgcaactacatgcgcgaagga 926
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 AACTGCACACGGGGGAAACACCTTTAAAGTGTACCTCTGCAACTACGATGCCAAAGGA 371

Qy 927 gagatgcctcacgggacacaccttaggacacattctg 962
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GAGATGCGCTCACGGGACACCTTAGGACACATTCTG 335

RESULT 9
AA866707/c 936 bp mRNA EST 16-MAR-1998
LOCUS vx85a10.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1281978 5'
DEFINITION similar to gb:LQ3547 Mouse ikaros DNA binding protein (MOUSE);,
      mRNA sequence.
ACCESSION AA866707
NID 92862152
VERSION AA866707.1 GI:2962152
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 936)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
COMMENT On Jan 17, 1998 this sequence version replaced gi:2044537.

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:673778
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 499.
Location/Qualifiers
1..936
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1281978"
/clone_lib="Soares 2NbMT"

FEATURES
SOURCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 217)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I- oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."
PASE COUNT      212 a   241 c   267 g   216 t
ORIGIN

Query Match      8.2%; Score 161.8; DB 39; Length 936;
Best Local Similarity 62.2%; Pred. No. 5.1e-29;
Matches 289; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

Qy 1436 ccgcagagatggaaaagaaacgcatcctctgcagagaaagatcttgcctctgaacga 1495
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 CAGGACGCCGTGGATTAACCTGCTGCTGTCTCCCAAGCCAAAGTCTGTGTCATGGAGCGA 679

Qy 1496 ggtctgtcccccaataacagtgcccaggaactccacagacacccagacagacacagagat 1555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 GAGCCTTCCCGGACCAACAGCTGCCAAGACTCCACAGATACAGAGAGCAACCGCGAGAAC 619

Qy 1556 cgc-caacatctctaccagcaaaagccagctggtctctccccccagccgcaatggatgcc 1614
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 AGCCACAGCGGCTTATTCTTACCTACCAACCAACATCAACCCGATGCACGAATGGGTGGC 559

Qy 1615 tctctgaagagatgctcctcctctcttttgaaactctcaagccctccctccatctgcctgag 1674
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 ---TCTCAAGGAGGAGGAGCGCGCTTACGAGGTGCTGAGGGCGGCTCAGAGAACTCGCA 502

Qy 1675 ggactccatcaaatgatcatcaaaagaggaggatggtggtggtggtggtggtggtggtggtg 1734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 GGATGCTTCCGTGTGGTGCAGCAGAGTGGCGAGCGAGTGAAGGTGTACAAGTGCAGACA 442

Qy 1735 ctgcagctctctctctctctctctctctctctctctctctctctctctctctctctctctct 1794
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 CTGCGCGGTGCTCTCTCTCTGATCATCGTATGATACCATTCACATGGCTGCCATGGCTT 382

Qy 1795 ccgtgatccctttgagtgtaacatgtgtggctatcgaaagccacacatcgctatgattctc 1854
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 TCGGATCCCTTTGAGTGTAAATGTGTGGTTATCATCAGCCAGGACAGGTACGAGTTCTC 322

Qy 1855 ctctcacatcgccagagagagacagagccatgttgaagtgcagc 1899
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 ATCCCATATACGCGGGGAGCATCGTTACCACCTGAGCTAAAC 277

RESULT 10
AI040940/c 217 bp mRNA EST 27-AUG-1998
LOCUS ov53b04.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1641007
DEFINITION 3', similar to RF:008900 008900 AIOLOS ;, mRNA sequence.
ACCESSION AI040940
NID 93280134
VERSION AI040940.1 GI:3280134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 217)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```


Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:458077
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 485.

FEATURES

Location/Qualifiers
 1..517
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:749093"
 /clone_lib="Soares mouse 3NDMS"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TTTTACCAATCTGAAGTGGAGCGCGCTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M.Fatima Bonaldo."
 127 a 126 c 119 g 145 t

BASE COUNT
 ORIGIN

Query Match 14.1%; Score 279; DB 31; Length 517;
 Best Local Similarity 100.0%; Pred. No. 2.3e-57;
 Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1700 gaagggaggtgatggtgttcgatgtgaccatgccacgtcctctctctagattat 1759
 Db 1 GAAGGGAGGTGATGATGTGTTTCGATGTGACCATGCGACGTCTCTCTCTAGATTAT 60

QY 1760 gtgatgttcaccatccacatggggtgccatgggttcctgtgacccctttgagtgtaacatg 1819
 Db 61 GTGATGTTCCACCATCCACATGGGTGCCATGGTTCCTGTGATGATCCCTTTGAGTGAACATG 120

QY 1820 tttggtctatcgaagcagcagatcgctatgagttctctctcaccatcgccagagagagcac 1879
 Db 121 TGTGGCTATCGAAGCCACGATCGCTATGAGTTCTCTCTCACATCGCCAGGAGAGCAC 180

QY 1880 agagccatgttgaagtgaagcagcatctctcctcaatgcgaggggtcaacatgttttttaaacg 1939
 Db 181 AGAGCCATGTTGAAGTGAGCATCTCTCTCAATGCGAGGGTCAACATGTGTTTTTAAAGC 240

QY 1940 tgatgttagccttatccagtagactgaactcaaacccac 1978
 Db 241 TGATGTAGCCTTATCCAGTAGACTGAACCTCAAAACCCAC 279

RESULT 6
 AI644775

LOCUS
 DEFINITION
 vbl6d03.y1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:749093
 5' similar to gb:L03547 Mouse Ikaros DNA binding protein (MOUSE);,
 mRNA sequence.

ACCESSION
 AI644775
 NID
 VERSION
 AI644775.1 GI:4723250
 KEYWORDS
 EST.
 SOURCE
 house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 401)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,

Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
 Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 On May 7, 1998 this sequence version replaced gi:3121606.

TITLE
 JOURNAL
 COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:458077

This read is a RESEQUENCE of a previously sequenced mouse clone
 correct orientation)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 358.

FEATURES
 source

Location/Qualifiers
 1..401
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /map="2"
 /clone="IMAGE:749093"
 /clone_lib="Soares mouse 3NDMS"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCGCTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M.Fatima Bonaldo."
 98 a 100 c 90 g 113 t

BASE COUNT
 ORIGIN

Query Match 14.1%; Score 279; DB 49; Length 401;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1700 gaagggaggtgatggtgttcgatgtgaccatgccacgtcctctctctagattat 1759
 Db 1 GAAGGGAGGTGATGATGTGTTTCGATGTGACCATGCGACGTCTCTCTCTAGATTAT 60

QY 1760 gtgatgttcaccatccacatggggtgccatgggttcctgtgacccctttgagtgtaacatg 1819
 Db 61 GTGATGTTCCACCATCCACATGGGTGCCATGGTTCCTGTGATGATCCCTTTGAGTGAACATG 120

QY 1820 tttggtctatcgaagcagcagatcgctatgagttctctcaccatcgccagagagagcac 1879
 Db 121 TGTGGCTATCGAAGCCACGATCGCTATGAGTTCTCTCTCACATCGCCAGGAGAGCAC 180

QY 1880 agagccatgttgaagtgaagcagcatctctcctcaatgcgaggggtcaacatgttttttaaacg 1939
 Db 181 AGAGCCATGTTGAAGTGAGCATCTCTCTCAATGCGAGGGTCAACATGTGTTTTTAAAGC 240

QY 1940 tgatgttagccttatccagtagactgaactcaaacccac 1978
 Db 241 TGATGTAGCCTTATCCAGTAGACTGAACCTCAAAACCCAC 279

RESULT 7

```

|||||
Db 301 GAAATAGCGCACTATGAACGATTAAGTGTGAGAGACAGTGCCTTATGACACACAGA 360
QY 707 ccaacagtgagggaagtgagctgcagctgtgctggtgtatctctgcattagcttcaacgtc 766
|||||
Db 361 CCAACCGGTGGGAAGATGAATCGGAGCTGTGCGGGTTATCTTGCATTAGCTTCAACGTC 420
QY 767 ttgatgttcataagcgaagccataccgcgaacgcccgttcagtgtaatacagtcgagg 826
|||||
Db 421 TNGATGTTTCATAAGCGAAGCCATACCCGGAACCCGCGTCCAGTGGTATCATGTCGCG 480
QY 827 gcatcttttactcagaaggt 847
|||||
Db 481 GCATTCTTTACTCAGAAAGT 501

RESULT 4
AI509603/C
LOCUS
DEFINITION
AI509603 480 bp mRNA EST 12-MAR-1999
vx13e05.y1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1264352 5'
similar to gb:L03547 Mouse Ikaros DNA binding protein (MOUSE);,
mRNA sequence.
ACCESSION AI509603
NID 94408508
VERSION AI509603.1 GI:4408508
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 480)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Perron,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188293.
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:666904
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 467.
Location/Qualifiers
1. 480
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1264352"
/clone_lib="Soares 2NbMT"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACCATCTGAAGTGGAGCGCGCGGTCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two

```

```

rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 88 a 112 c 159 g 121 t
ORIGIN
Query Match 23.8%; Score 472; DB 47; Length 480;
Best Local Similarity 99.0%; Pred. No. 7.2e-104;
Matches 475; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1183 tcagaaattcatcgatgagaaagcgcactgcttcagtcgcaactacaatcccggtacat 1242
Db 480 TCAGAAATTCATCGGTGAGAACGGCACTGCTTCGATGCCAACTACATCCCGGCTACAT 421
QY 1243 gtacgagaaggagaacgagatgatgcagaccgggatggaccagccatcaataaacgc 1302
Db 420 GTACGAGAAGGAGAAGCAGATGATGCAGACCCGGATGATGGACCAAGCCATCAATAACGC 361
QY 1303 catcgactatctaggggtgaagcccttcgccccttagtcagactccgctgctccac 1362
Db 360 CATCACTATCTAGGGGCTGAAGCCCTTCGCCCTTAGTCCAGACTCCGCCCTGCTCCAC 301
QY 1363 ctctgagatggtcccgatcatcagcagtggtaccccatagcacttactcgggcccagat 1422
Db 300 CTCTGAGATGTTCCCACTCATCAGCAGTGTGTACCCCATAGCAGCTTACTCGGGCCGATAT 241
QY 1423 gccaatggggcccccgcaggagatggaagaaacggatcctcctgccagagaaagatctt 1482
Db 240 GCCCAATGGGGCCCGCAGGAGATGGAAGAAACGGATCCTCTCTGCCAGAGAAGATCTT 181
QY 1483 gctctgaacagagtgctgtccccaataaagtgccaggactccacagacacgacag 1542
Db 180 GCTTCTGAACAGAGTCTGTCTCCCAATAACAGTCCCGAGACTCCACAGACACGACAG 121
QY 1543 caaccacagagatgcccaatctctaccagcaagcagctgtctctcccccagccgcg 1602
Db 120 CAACACAGGAGTATGCCCAATCTCTACAGCAAGACACAGTGTCTCTCCCGCCGCG 61
QY 1603 caatggatgctcttctgaaggagtgctctcttttgaactctctaaagcccccctcc 1662
Db 60 CAATGGATGCTCTTCTGAAGGAGTCCCTCGCTCTTTTGAACCTCTCAAGCCCCCTCC 1
RESULT 5
AA290536 517 bp mRNA EST 14-APR-1997
LOCUS vb16d03.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:749093
DEFINITION 5' similar to gb:L03547 Mouse Ikaros DNA binding protein (MOUSE);,
mRNA sequence.
ACCESSION AA290536
NID 91937175
VERSION AA290536.1 GI:1937175
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 517)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393523.
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Published (1999)

BASE COUNT	112 a	150 c	149 g	162 t	4 others
ORIGIN					

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

782	Qy	782	cgaagccataccggcgcaacgcccggttcacgtgtaatacagtcgaggcgcatcttttactcag	841
577	Db	577	CGAAAGCATTTACCGCGAACGCCCGTCCAGTGTAAATCAGT-TGGGGCATCTTTTACTCGN	519
842	Qy	842	aaaggtaaacctcctcgtcatattaaactgcacacgggggaaaaaaccttttaagtgtcac	901
518	Db	518	AAGGTTACCCNTCCNTGTCATATTATTAACCTGCACACGGGGGAAAAACCTTTTAAAGTGTAC	459
902	Qy	902	ctctgcaactacgcatacgccaaaggagagatgcgctcacgggacacc-ttaggcacacattc	960
458	Db	458	TNCTGCAACTACGCATGCCAAAGGAGAGATGCGCTCAGGGACACCTTTAGGCACATTC	399
961	Qy	961	tgtggggaagccgtacaaagtgtgagttctcggggaagctacaagcagagaagctccct	1020
398	Db	398	TGTGAGAAAGCCGTACAAAGTGTGAGTTCTCGGGAAGAGCTACAAAGCAGAGAAGCTCCCT	339
1021	Qy	1021	ggaggagcacaaggaagcgtgcgagcttttcttcagaacccctgacctggggagcgtgc	1080
338	Db	338	GGAGGAGCACAAAGAACGCTGCCGAGCTTTTCTTCAAGACCCTGACCTGGGGGACGGTGC	279
1081	Qy	1081	aagtgtgagggcaagcacacatcaaaagccgagatgggaagtgaagagagctctcgtccctgga	1140
278	Db	278	AAGTGTGGAGCGAAGACACATCAAAAGCCGAGATGGGAAGTGAGAGAGCTCTCGTCTCGGA	219
1141	Qy	1141	cagattgaagcgaatgctgctaaagcgaagctcgtgcctcagaataatcatcgtgtga	1200
218	Db	218	CAGATTASCAAGCAATGTGGCTTAAGCAAGAAAAGCTCGATGCTCAGAAAATTCATCGGTGA	159
1201	Qy	1201	gaagcgcactgcttcgatgccaaactacaatcccgctacatgtacagaagaggagaacga	1260
158	Db	158	GAGCGGCATCTGCTTCGATGCCAACTACAAATCCCGGCTACATGTACCGAAGAGGAGAACGA	99
1261	Qy	1261	gatgatgcagaccgggatgatggaccaagccatacaataacgccctcagctatctcaggggc	1320
98	Db	98	GATGATGCACACCCGGATGATGGACCAAGCCATCAATAACGCCATCAGCTATCTA-GGGC	40
1321	Qy	1321	tgaagccttcgcgcccttagtccagatcgcctgcctcc	1359
39	Db	39	TGAAGCCCTTCGGCCCTTAGTCCACATCCGCCCTGCCTC	1

AI550402	AI550402	501 bp	mRNA	EST	23-MAR-1999
LOCUS	vx13e05.x1	Soares 2NbMt	Mus musculus	cdna clone	IMAGE:1264352 3'
DEFINITION	similar to TR:008900 008900 AIOLOS ; mRNA sequence.				

similar to TR:008900 O

ACCESSION	AI550402
NID	94482765
VERSION	AI550402.1
KEYWORDS	GI:4482765 EST.

SOURCE	house mouse.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 501)

BASE COUNT	158 a	118 c	143 g	81 t	1 others
ORIGIN	SOARES and M.FACIMA BONARDU.				

Query Match 24.4%; Score 484; DB 48; Length 501;
Best Local Similarity 97.8%; Pred. No. 9.5e-107;

347	gctgcagcccgcgcgccgcgacatggaagatatcaaacgcactgaggactgaaa	406
Qy		
Db		
1	GCTGCAGCCCGCGCGCCGCGACATGAAGATATACAACGCAGCTGTGGAGCTGAAA	60
Qy		
407	agcagggagagcagcctctgtccacagagagcccgacgcctctgaatgactacagcttg	466
Qy		
61	AGCAGGAGGAGCAGCCTCTGTCCACAGAGAGCCGACGGCTCTGAATGACTACAGCTTG	120
Db		

QY	467	cccaaacctcatgagatagataaaacgctggacagtagagaagcccccagcccaatgaaagacgaa	526
Db	121	CCCAAACTCATGAGATAGAAAACGTGGACAGTAGAAAGCCCGCAGCCAATGAAGACGAA	180
QY	527	gatcagggagaagattcgatgaaagtgaagatgaatacacagcgacagagatgagaacatt	586
Db	181	GATCAGGGAGAAGATT CGATGAAAGT GAAAGATGAATACAGCGACAGAGATGAGAAACATT	240
QY	587	atgaagccgaagcccatggaagatgcagagaagagtgaaatgcccttaccgctatgcgaaga	646
Db	241	ATGAAGCCGGAGCCCATGGGAGATGCGAGAGAGAGTGAATATGCCCTTACAGCTATGCRAAGA	300
QY	647	gaatacacgcgaactatgaagacattaaactgaagagacacatgccctatgacaaacagcaga	706

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 08:08:27 ; Search time 437.25 Seconds
(without alignments)
8950.272 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagagcgacacgcgtcgg.....gaactcaaacccacctcgag 1984

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database : EST:

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	543.4	27.4	546	48	AI549961
2	503.8	25.4	577	39	AA863970
3	484	24.4	501	48	AI550402
4	472	23.8	480	47	AI509603
5	279	14.1	517	31	AA290536
6	279	14.1	401	49	AI644775
7	269	13.6	332	39	AA867722
8	204	10.3	609	40	AA920812
9	161.8	8.2	936	39	AA866707
10	154.6	7.8	217	41	AI040940
11	113.6	5.7	557	45	AI325349
12	87.2	4.4	479	46	AI428749
13	86.8	4.4	339	38	AA808413
14	85.2	4.3	355	24	H83408
15	84	4.2	356	38	AA764246
16	81.4	4.1	389	28	AA119538
17	81.4	4.1	576	31	AA277407
18	80.2	4.0	382	33	AA400326
19	80	4.0	494	39	AA848430
20	79.2	4.0	424	26	W19169
21	79	4.0	509	27	AA027561
22	79	4.0	271	47	AI511308
23	78.8	4.0	662	26	W28371
24	78.4	4.0	685	26	W27601
25	77.4	3.9	550	27	AA014148
26	76.6	3.9	805	49	AI648952
27	76.2	3.8	552	31	AA289328
28	76	3.8	245	20	Z21189
29	76	3.8	577	43	AI179640
30	75.4	3.8	470	21	RI7304
31	75.4	3.8	467	21	R19256
32	75.4	3.8	431	22	R22824
33	74.8	3.8	512	46	AI06535
34	74.4	3.8	394	27	AA003426
35	74.4	3.8	438	33	AA427841
36	74.4	3.8	485	33	AA451439
37	74.2	3.7	505	29	AA144904
38	74.2	3.7	406	54	HSN009139
39	74	3.7	309	24	H73895
40	73.8	3.7	644	22	H05551
41	73.6	3.7	601	34	AA455712
42	73.6	3.7	497	34	AA499468
43	73.6	3.7	484	42	AI124541
44	73.4	3.7	496	40	AA920944
45	73.2	3.7	392	23	H28944

ALIGNMENTS

RESULT	1	AI549961	546 bp	nrna	EST
AI549961	AI549961	AI549961	546 bp	nrna	23-MAR-1999
LOCUS	vx02e08.x1	Soares 2NbMT Mus musculus cdna clone IMAGE:1263302	3'		
DEFINITION	similar to TR:008900	008900 AIOLOS ;, mrna sequence.			
ACCESSION	AI549961				
NID	94482324				
VERSION	AI549961.1	GI:4482324			


```

; MOLECULE TYPE:  CDNA
; FEATURE:
; NAME/KEY:  CDS
; LOCATION:  1..1170
PCT-US95-09345-5

```

Query Match	11.98;	Score 235.2;	DB 5;	Length 1170;
Best Local Similarity	57.1%;	Pred. No. 7e-54;		
Matches 595;	Conservative 0;	Mismatches 393;	Indels 54;	Gaps 7;
961	atattaaactgcacacgggggaaacaccttttaagtgtcacctctgcaactacgcatgcc	920		
143	AGAGTGCATCGAGGCATGGGTGAACGGGCTTTCAGTGCACACCACTGTGGGGCTCCTTTA	202		
921	aaagagagatgcgtcacggggacacctttaggacacattctgtggagaagcogtacaagt	980		
203	CCAGAAAGGCAACCTCTCGGGACACATCAAGCTGCATCTGGGTGAAGAAGCCCTTCAAA	262		
981	gtgagtctgcgggaagaagctacaagcagagaagctccctggaggagcacaagaacgct	1040		
263	GCCATCTTTGCAACTATGCCCTGCCCGAGGAGCGCCCTACCGGCCACTGAGAGCGC	322		
1041	gccgagcttttttcagaaacctgacctggggagcgtgcgaagtgtggaggcaagacaca	1100		
323	ACTCGGCTATTAAAGGAAGAACTAACCAACAGAGATGGC-----AGAG	367		
1101	tcaagcgcgagatggaaagtgcagagagctctcgtctctggacagattagaagaacatgtgg	1160		
368	ACCTGTGC AAGATAGGACGAGAGAGTCCCTGTCTTGGACAGGCTTGGCAAGCAATGTGC	427		
1161	ctaagcgaaaaagctcgtgctcagaaattcatcggtgagaagcgcgactgtcttcgatg	1220		
428	CCAAACGTAAGAGCTCTATGCTCAGAAATTTCTTGAGACAAAGTGCCTGTCAAGATGC	487		
1221	ccaactacaatccccggtctacatgtacagagaaggagaacgagatgacgagaccggatga	1280		
488	CTATGACAGTGCC-----AACTATGAGAAGGAGGATATGATGACATCCACGCTGA	538		
1281	tggaccaagcacaataaacgccatcagctatctaggggctgaagccttcggccctctag	1340		
539	TGGACGAGGCCATCAACATGCCATCACTTACCTTGGGGGCTGAGTCCCTGGCCCATTTGG	598		
1341	tccagactcgcgctctcccacctcttgatgtgtgtccacgtcatcagcagttgtacccca	1400		
599	TGCAGACACCCCGCGGT---AGCTCCGAGGTGGTGCAGTCAATGATGATGATGATGATG	655		
1401	tagcacttactcggcgccgatgccaatggg-----gcccgacgagaatgg	1468		
656	TGCACAGCCCCCTCAGATGGCCCCCAGCGTCCACCAATTGACGACAGAGCGCGGTGG	715		
1449	aaagaacacggatcctctgcagagaagatcttgccttctgaacgaggtctgtcccca	1508		
716	ATAACTTGCTGTGCTGTCCAAAGCCAAAGTGTGTATCGGACGAGAGCGGCTCCCGA	775		
1509	ataacagtgccagagactccacagaccgacagacacccacgagg---atcgccaacatc	1565		
776	GCAACAGCTGCCCAAGACTCCACAGATACAGAGACACACGCGGAGGAAACAGCGCAGCGG	835		
1566	tctaccagcaaaagccaagtgtctctcccacaggcccgcaatggatgctctctgaagg	1625		
836	TTATCTACTTAACCAACACATCAACCCGCATGACGCAATGGGCTGGC---TCTCAAGG	892		
1626	aggtccctcgtctcttttgaaactctcgaagccccctcccatctcgtgagggactccatca	1685		
893	AGGACAGCGCGCCCTTACGAGGTGTGTAGGGCGGCGCTCAGAGAACTCGCAGGATGCTTCC	952		
1686	aaagtatacaaaaagggagagtgcagatgtgttttcgatgtgaccactgcacagctcc	1745		
953	GTGTGCTACGACAGAGTGGCAGAGCAGCTGAAGGTTGTCAAGTGGCAACACTGCGCGGTGC	1012		
1746	tcttcttagattatgtgtgttccacctccacatg-----gggtgcattgggtttcc	1796		

```
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1170
US-08-485-590-6

Query Match 11.9%; Score 235.2; DB 3; Length 1170;
Best Local Similarity 57.1%; Pred. No. 7e-54;
Matches 595; Conservative 0; Mismatches 393; Indels 54; Gaps 7;

Qy 861 atattaaactgcacacgaggggaaacatttttaagtgcacactctgcacactacgcatgcc 920
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 143 AGAGTGATCGAGGATGGGTGAACGGCTTTCCAGTGAACCAAGCTCTGGGGCTCTTTA 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 921 aaagagagatgcgtcacggacacaccttagacacattctgtgagaaagcgtacaagt 980
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 CCCAAAGGCAACCTCTCTGGCGACATCAAGCTGCATCGGGTGAAGAGCCCTTCAAT 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 981 gtgattcttcggaagaactcaagcagagagtgcttccttgagagcacaaggaacgct 1040
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 263 GCCATCTTGAACATATGCTCCCGCGGAGGAGCGCCCTCACCGGCCACCTGAGAGCG 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1041 gccagcgttttttcagaaacctgcacctggggagcgtgcgaagtgtgagcaagacaca 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 323 ACTCGTGCATTAAAGGAAGAACTAACCAACAGAGATGCG-----AGAAG 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1101 tcaagcagagatggagtgagagagctctgctctgagacattagcaagcaatggg 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 ACCTGTGAAGATAGAGCAGAGAGGTCCCTTGTCTGGACAGGCTGGCAAGCAATGCG 427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1161 ctaagcgaagagctcgtgcctcagaaattcctcgtgagagcggcactgctcgtatg 1220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 CCAACGTAAGAGCTCTATGCTCAGAAATTTCTTGAGACAGAGTGCCTGTGACAGATGC 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1221 ccaactacaatcccgggtacatgtacgagaaggagagagatgatgcagaccocggatga 1280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 CCTATGACAGTGCC-----AACTATGAGAAGGAGGATATGATGACATCCACAGTGA 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1281 tggaccaagccataataacgcatcagctatctaggggctgaagccttcggcccttag 1340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 TGGACAGGCCCAACCAACATGCATCAACTACCTGGGGGCTGAGTCCCTGGCCCATGG 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1341 tccagactccgctcctcccactctgcagatggtcccagtcacacagcagtggtaccacca 1400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 TGCACACACCCCGGT---ACGTCGAGGTGGTGCCAGTCTATCATGCTCCATGTACCAGC 655
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1401 tagcacttactcgggcccgcataatggggt-----gccccgcaggagatgg 1448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 TGCACAAGCCCGCTCAGATGCGCCCGCCACGGTCCCAACCATTCAGCAGAGGACGCGGTG 715
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1449 aaaaagaacggatctctctgcagagaagatcttccctctgaaagaggtgtgtcccca 1508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 716 ATAACTTGTGCTGTGTCGAAGGCCAAAGTCTGTGTCATCGGAGCGGAGAGCCCTCCCGA 775
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1509 ataacagtgcacagactccacagacacgcagacagcaaacacagagg---atgcaccaatc 1565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 776 GCAACAGCTGCCAAGACTCCACAGATACAGAGAGGCAACGCGGAGGAGCAACGCGCGCC 835
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 12
PCT-US95-09345-5
; Sequence 5, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```


RESULT 9
US-08-465-590-8
; Sequence 8, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,590
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1002
US-08-465-590-8

Query Match 13.6%; Score 269.6; DB 3; Length 1004;
Best Local Similarity 58.9%; Pred. No. 4.1e-63;
Matches 602; Conservative 0; Mismatches 369; Indels 51; Gaps 6;

Qy 878 ggggaaaaacccctttaagtgtcacctctgcaactacgcatgcaaaaggagagatgcgctc 937
Db 1 GGAGAACGCCCTTCAGTGCATCAGTGGGGGCTCATTCACCCAGAAAGGCCAACCTG 60
Qy 938 acgggacaccccttaggacacattctgtggagaagccgtacaagtgtgagttctgcggaaga 997
Db 61 CTCGGCACATCAAGCTGATCCGGGAGAACCCCTTCAAAATGCCACCTCTGCAACTAC 120
Qy 998 agctcaagcagagaagctccctggaggagcacaaggaacgtgcgagcttttcttcag 1057
Db 121 GCTCCCGCGGAGGACGCCCTCACTGGCCACCTGAGGACGCACCTCCGCTCATTAAGAA 180
Qy 1058 aaccctgacctggggacgctcaagtgtggaggcaacacatcaaaagccgagatgga 1117
Db 181 GAACTAGCACAGTGAATGCG-----AGAACTGTGCAAGATAGGA 225

Qy 1118 agtgagagagctctcgtctctgacagattagcaagcaatgtggctaaagcaaaagctcg 1177
Db 226 TCAGAGAGATCTCGTGTGGACAGACTAGCAAGTAATGTGCGCAACGTAAGAGCTCT 285
Qy 1178 atgacctcaaaattcatcgtgtgagaagcgcactgtcttcgatgccaatacaatccccgc 1237
Db 286 ATGGCTCAGAAATTTCTTTGGGACAAGGGCCTGTCCCGACACGCCCTACGACAGTGC--- 342
Qy 1238 tacatgtacgagaaggaacagatgatgcagaccggatgatggaccagagccatcaat 1297
Db 342 --CAGGTACGAGAGGAGAACGAATGATGAAGTCCACAGTGTGAGCAGCAAGCCATCAAC 399
Qy 1298 aacgccatcagctatctaggggctgaagccttcgcgccttagtccagactccgcctgct 1357
Db 400 AACGCCATCAACTACCTGGGGCGAGTCCCTGCGCGCGTGGTGCAGACGCCCGCG-- 458
Qy 1358 ccacacctgagatgggtcccsagtcatacagcagtggtgtaccat-----agca 1405
Db 458 -GCGGTTCCGAGGTGGTCCCGGTATCAGCCCGATGTACACCGCTGCACAGCGCTCGGAG 516
Qy 1406 ctactcgggcccgatatgccaatggggcccgagagatggaaaagaaacgagatcctc 1465
Db 517 GGCACCCCGCGCTCCACCACTCGGCCAGACAGCGCGGTGGAGTACTCTGCTGCTC 576
Qy 1466 ctgcagagaagatcttgccttctgaacgagtgctgtcccccataaacagtgcacagac 1525
Db 577 TCCAAGGCCAAGTTGGTGCCTCGGAGCGGAGCGCTCCCGAGCAACAGCTGCCAAGAC 636
Qy 1526 tccacagacaccgacagcaaccacagagatc-----gccaaatctctaccagcaagc 1579
Db 637 TCCACGGACACCGAGAGCAACACAGGAGGACGCGCGTCTTATCTACTTGACCAAC 696
Qy 1580 cagtggtctcccccagggccgcaatgggagtgctcttctgaagaggtccctcgctct 1639
Db 697 CATATCGCCCGCGCGCAACGCGGTCTG-----CTCAGGAGGAGGACGCCCGCC 747
Qy 1640 ttgaactctcctcaagccccctccatctgctgagggagctccatcaaaagtatcaacaa 1699
Db 748 TAGCACTGCTGCGCGCGCGCTCGGAGAACTCGCAGGACGCGCTCCGCGTGGTCAAGACC 807
Qy 1700 gaaggggaggtgatgtgttcgatgtgaccactgcacgctcctcttctagattat 1759
Db 808 AGCGGGAGGACAGATGAAGGTGTACAAGTGCAGAACACTGCCGGGTCTCTTCTGGATCAC 867
Qy 1760 gtgatgttcaccatccacatgggtgccatgttccgtgatccctttgagtgtaacatg 1819
Db 868 GTCATGTACACCATCCACATGGGTGCGCACCGCTTCCGTGATCTCTTTGAGTGCACATG 927
Qy 1820 tgtggctatcgaagccacgcatcgctatgagttctctctctcacatgcgcagaggagacac 1879
Db 928 TCGGGCTACACAGCAGGACCGGTACGAGTTCTCGTGCACATAACCGGAGGGGAGCAC 987
Qy 1880 ag 1881
Db 988 CG 989
RESULT 10
PCT-US95-09345-7
; Sequence 7, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:

QY 1276 gatgagggagccatcaataacgccatagctatctaggggtgaagccttcgccc 1335
|||||
Db 1140 CGTGATGGACCACTCAACAAATGCCATACCTGGGGCTGAGTCCCTGGCC 1199
QY 1336 cttagccagactccgctgctccaccctagatgggtccagtcacagcagtgta 1395
|||||
Db 1200 ATTGGTGAGACACCCCGG---AGCTCCGAGGTGGTCCAGTATCAGCTCCATGTA 1256
QY 1396 cccatagcacttactcgggcccagatgccaatggggccc-----ccgcagga 1443
|||||
Db 1257 CAGCTGCACAAGCCCTTCAGATGGCCGCCACCGTCCACCAATTCAGCAGGACGC 1316
QY 1444 gatgaaaagaacgatactctcgcagagagatcttgccttgaacagggtctg 1503
Db 1317 CGTGGATACTTGTCTGCTGTGCTCAAGCCAAAGTCTGTCTCATGGAGGAGCCCTC 1376
QY 1504 ccccaataacagtcgagagactccagacccagacagacacacagag---atgccca 1560
|||||
Db 1377 CCGAGCACAGCTGCCAGACTCCACATACAGAGCAACCGCGAGGAGACAGCGCAG 1436
QY 1561 acatctcaccagcaagcagctggtctccccagggccgcaatgggagctcttct 1620
|||||
Db 1437 CGGCTTATCTACTTAACCAACACATCAACCGCATGCACGCAATGGCTGCTGCT 1493
QY 1621 gaaggaggctcctctcttcttgaactctcgaagccctccctcctcctcctcctc 1680
|||||
Db 1494 CAAGGAGGAGCAGCGGCTACAGAGTGTGAGGGCGGCTCAGAGAACTCGCAGATGC 1553
QY 1681 catcaagtgatacaacaagaagggaggtgatggatgttctcagatgacacagca 1740
Db 1554 CTTCCTGTGGTTCAGCAGCAGTGGCAGGAGCTGAAGGTGTACAAGTGGCAACACTGCGC 1613
QY 1741 cgtctctctcctagattatgtgtttcaccatccacatg-----gggtggcattgg 1791
Db 1614 CGTCTCTCTCTGATCAGCTCATGTATACCATTCATCATGCTGGGTGCCATGGTCCATGG 1673
QY 1792 ttccgtgaccccttggatgtaaacatgtgtggtctatcgaagccagatcgctatagtt 1851
Db 1674 CTTCGGGATCCCTTGGTGTAAATGTGTGGTTTATCAGCAGGAGCAGTACGAGTT 1733
QY 1852 ctctctcacatcgccagagagacacagagccatgttgaagtggc 1899
Db 1734 CTATCCCATATCACCGGGGGAGCATCGTTACCACCTGAGCTAAAC 1781

RESULT 6
US-08-465-590-2
Sequence 2, Application US/08465590
Patent No. 5824770
GENERAL INFORMATION:
APPLICANT: Georgopoulos, Katia A.
TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
NUMBER OF SEQUENCES: 164
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MPG-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1788 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 223..1515
US-08-465-590-2

Query Match 17.5%; Score 346.4; DB 3; Length 1788;
Best Local Similarity 62.0%; Pred. No. 1.5e-83;
Matches 715; Conservative 0; Mismatches 376; Indels 63; Gaps 8;

QY 794 ggcgaacccctccagtgtaatacagtcggtggggcatcttttactcagaagaagtaacctc 853
|||||
Db 382 GGTGAACGGCCCTTCCAGTGCACACAGCTGTGGGGCTCTTTACCAGAAAGGCAACCTC 441
QY 854 ctccgtcattataactgcacacgggggaaaccttttaagtgtcacctctgcaactac 913
Db 442 CTGCGGCACATCAAGTCACCTCGGTGAGAGCCCTTCAATGCCATCTTTGCAACTAT 501
QY 914 gcatgccaaagagagatgctcagcgggacacaccttaggacacattctgtggaagcgc 973
Db 502 GCCTCGCGCGGAGGAGACGCCCTCACCGGCCACCTGAGGACGACCTCCCTTGTAAAGCT 561
QY 974 tacaagtgtgacttctcgggaagaagctacaagagagaagctccctgagagacaaag 1033
Db 562 CACAATGTGGATATTGTGGCCGGAGCTATAACAGCAGGCTCTTTAGAGGAGCATAAA 621
QY 1034 gaacgtgcccagagcttttctcagaacccctgacctgggggacg-----ctgcaagtgtg 1087
Db 622 GAGCGATGCCACAACCTACTTGGAAAGCATGGGCCCTTCGGGGCTGTGCCAGTCATTAAG 681
QY 1088 gaggcaagacacatcaag-----ccgagatgggaagtgaagagact 1129
Db 682 GAAGAAGCTTAACCAACAGAGATGGCAGAGACCTGTGCAAGATAGGAGCAGAGAGTCC 741
QY 1130 ctgcctcctggacagattagcaagcaatgtgctaagcaaaaagctcgatgctcagaaa 1189
Db 742 CTGTCTCTGGACAGGCTGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCTCTCAGAAA 801
QY 1190 ttcatcgttgagaagcggcactgcttcgatgccaactacaatccccgggtcacatgacag 1249
Db 802 TTTCTTGGAGACAAGTGCCCTGTGACACATGCCCCATGACAGTGCCAACCTATGAGAAGGAG 861
QY 1250 aagggaacgagatgatgcagaccccggaatgagaccagccatcaataacccatcagc 1309
Db 862 -----GATATGATGACATCCACGCTGATGGAGCAGGAGCCATCAACAATGCCATCAAC 912
QY 1310 tatctagggctgaagccttcgccccttagtccagactccctgctccaccctcctag 1369
Db 913 TACCTGGGGCTGAGTCCCTGCGCCCAATTGGTGACAGACACCCCCCGGT---AGCTCCGAG 969
QY 1370 atggtccagctcatcagcagtggtaccaccaatgacacttactcggggccgagatgccaatg 1429
Db 970 GTGGTGCCAGTCATCAGCTCCATGTACCAGCTGCACAGCCCCCTCAGATGGCCCCCA 1029

Qy 1216 cgaatgcaactacacccggtacatgtacgagaagagacgagatgatgcagaccg 1275
Db 1089 CATCCCTATGACAGTGCCTAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139
Qy 1276 gatgatgaccacgacatcaataacgcatcagctatctaggggtgaagccttccgcc 1335
Db 1140 CGTGATGGACCGCCATCAACATGCATCACTACTCTGGGGCTGAGTCCCTGGCC 1199
Qy 1336 cttagtcagactccgctgtccacactctgagatggtccagctacatcagcagtgtga 1395
Db 1200 ATTGGTCGACAGCCGCCCGGT---AGTCCGAGGTGTGTCAGTCATCAGTCCATGTA 1256
Qy 1396 ccccatagcacttactcggcggtatgccaatggggccc-----ccgagga 1443
Db 1257 CCAGCTGCACAAGCCGCCCTCAGATGCGCCGCCACGCTCCCAACCATTCAGCACAGGACGC 1316
Qy 1444 gatgaaagaaacggtccctcctgcagagaagatcttgctcttgaagcaggtctgtc 1503
Db 1317 CGTGGATAACTTGTGCTGCTGCCAAGGCCAAGTCTGTGTCATCGGAGCGAGGCCCTC 1376
Qy 1504 ccccaataacagtgcccgagactccacagacccagcagcaaccacaggg---atcgcca 1560
Db 1377 CCCGAGCAACAGCTGCCAAGACTCCACAGATACAGAGACAAACGCGGAGGAACAGCGCAG 1436
Qy 1561 acatctctaccagcaagcagcgtggtcctcccccagcgcgcgaatggatgctcttct 1620
Db 1437 CGGCTTATCTACCTACCAACCAACATCAACCCGCAATGCAGCAATGGGTGGC---TCT 1493
Qy 1621 yaagaggtccctcgtcttttgaactcctcaagcccccctccatctgctgagggactc 1680
Db 1494 CAAGGAGGAGCAGCGCCCTACGAGGTCTTGAGGGCGCCCTCAGAGAACCTCGCAGGATGC 1553
Qy 1581 catcaagtgatcaacaagaagggggggtgagtggtgttgcagtgtgaccactgcca 1740
Db 1554 CTTCCGTTGTGTCACGACGAGTGGCGGAGCAGCTGAAGGTGTATCAAGTGCAGAACTGCCG 1613
Qy 1741 cgtcctctctagattatgattgtttcacatccacatg-----gggtccatgg 1791
Db 1614 CGTCTCTTCTGGATCAGGTCATGTATACCATTCATCATGGGTGCCATGGCTGCCATGG 1673
Qy 1792 ttccctgtagccctttgagtgtaacatgtgtggtatcgaagccacacgcatgctatgatt 1851
Db 1674 CTTTCGGGATCCCTTTGAGTGTAAATGTGTGTTATCAGACCCAGGACGAGGTACGAGTT 1733
Qy 1852 ctctctctacatcgcagagagagagcagagcagagccatgttgaagtgcac 1899
Db 1734 CTCATCCCATATCAGCGGGGGAGCATCGTTACCACCTGAGCTAAAC 1781

RESULT 5
PCT-US95-09345-4
; Sequence 4, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09345
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300

; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 223..1776
; PCT-US95-09345-4

Query Match 19.1%; Score 379.6; DB 5; Length 2049;
Best Local Similarity 61.8%; Pred. No. 2.1e-92;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;

Qy 700 cagcagacacacagcgggaagatgaactgcgacgctgctgctgcttattcctgcattgctt 759
Db 549 CATTGCATCTCTTACGGAAACTAAAGTGTGATATCTGTGGATCGTTTGTGATCGGGCC 608
Qy 760 caacgtcttctgattggttcataaagcgaagccataccgagcagcccttccagtgtaatca 819
Db 609 CAATGTCTCATGTTTACAAAAGAGTCACTACTGTGTAACGGCTTTTCCAGTGCAACCA 668
Qy 820 gtgcggggcattcttactcagaagaagtaacctcctcctcgtcattataaactgcacacgg 879
Db 669 GTCTGGGGCTCTCTTACCAGAAAGCAACCTCTCTCGGCACATCAAGCTGCACTCGG 728
Qy 880 ggaacacccctttaaagtgtcacctctgcaactacgcatgccaaagagagatgcgtcac 939
Db 729 TGAGAGCCCTTCAATGCCATCTTGTCAACTATGCTGCCGCGGAGGAGCGCCCTCAC 788
Qy 940 gggacaccccttagacacattctgtgagaagccgtacaaagtgtgagttcttgcggaagaag 999
Db 789 CGCCACCTCAGGAGCGACCTCCGTGTGTAAGCCTCACAATGTGGATATTGTGCCCGGAG 848
Qy 1000 ctacaagcagagaagctccctgagagagacaaagcagcagcgtcgcgagcttttcttcagaa 1059
Db 849 CTATAAACACGGAAGCTCTTTAGAGGAGCATAAAGAGCGATGCCACAACACTACTTGAAG 908
Qy 1060 cctgacctggggagc-----ctcaagtgtaggagcagacacatcaag----- 1107
Db 909 CATGGGCTTCCGGGGCTGTGCCAGTCATTAGGAAGAAACTAACCCACACAGATGGC 968
Qy 1107 -----ccgagatgggaagtgagagagctctcctctgacagattagcaagcaa 1155
Db 969 AGAAGACCTCTGTCAAGATAGGAGCAGAGAGGTCCCTGTCTTGACAGAGGTGCAAGCAA 1028
Qy 1156 tgggctaagcgaagcagctcgtatgctcagaattcatcgttgagaagcggcactgctt 1215
Db 1029 TGTGCGCAAAACGTAAGAGCTCTATGCTCAGAAATTTCTTTGGAGACAAGTGCCTGTGAGA 1088
Qy 1216 cgaatgcaactacaccccggtcacatgtacgagaagagagagagatgatgcagaccg 1275
Db 1089 CATGCCCTATGACAGTGCCTCAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139


```
QY 1153 caatgtggttaagcgaagaaagctcgatcgctcgaagaattcatcgggtgagagcggcactg 1212
Db 642 TAATGTCGCCAAACGTAAGAGCTTATGCTTCAGAAATTTCTTGGGACAAAGGCGCTGTC 701
QY 1213 ctctgagtgcaactacatacccggtctacatgtacgagaagaaacagagatgatgcagac 1272
Db 702 CGACAGCCCTAGCAGCTGC-----CAGGTACGAGAGGAAACGAATGATGAAGTC 755
QY 1273 ccggatgagtgaccagccatcaataacgcccacatcagctatctatgagggtgaaagccttcg 1332
Db 756 CCACGTGATGACCAAGCCATCAACAACGCCATCACTACCTTGGGGCCGAGTCCCTGCG 815
QY 1333 cccttagtcagagactcgcgtctcccaactctgagatggtcccaagtcacagagtg 1392
Db 816 CCGCGTGTGCAGACGCCCGCG---GCGGTTCCGAGGTGTCGCCGTATCAGCCCGAT 872
QY 1393 gtaccccaata-----gcacttactcggccgcatatgccaatggggcccccga 1440
Db 873 GTACCACTGCACAGGCGCTCGGAGGGCACCCCGCGCTCCAAACATCTCGGCCCCAGGACAG 932
QY 1441 ggagatggaaagaaacggatcctctgccagagaaagatcttgcttctgaacgaggtct 1500
Db 933 CGCGGTGAGTACTGCTGCTCTCCAAGGCCAAGTTGGTGCCTCGAGCGCGAGGC 992
QY 1501 gtccccaataacagtgccaggactccacagacacacagacagacacacagagatc----- 1557
Db 993 GTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAAACAGGAGCAGCG 1052
QY 1557 -gccaacatctaccagcaagcaacgtggtctctcccccagccgccaatggatgcc 1614
Db 1053 CAGCGGTCTTATCTGCTGACCAACACATCGCCCGACGCGCGCAACGCGGTGTCG----- 1108
QY 1615 tcttgaaagggtccctcgtctctttgaaactcctcaagccccctccactcgtcgtgag 1674
Db 1108 ----CTCAAGGAGGAGCACCGCGCTACGACCTGCTGCGCGCGCTCCGAGAACTCGCA 1163
QY 1675 ggactccataaagtgatacaaaaaaagggagggatgagatggtgttctcgatgtgacca 1734
Db 1164 GAGACGCTCCCGGTGGTGACGACACGCGGGGAGCAGATGAAGGTGTACAAAGTGCAGAAC 1223
QY 1735 ctgcaagctccttctagattatgtatgttcacacatccacatgggtggtccatggtt 1794
Db 1224 CTGCGGGTCTCTTCTGTGATCAGCTCATGTACACCATCCATGGGCTGCCACGGCTT 1283
QY 1795 ccdtgatccctttgagtgaacatgtgtggtctatcgaaagccacgacgtatgattctc 1854
Db 1284 CCGTGATCCTTTTGAAGTGCAACATGTCGGCTACCAACAGCCAGGACCGGTACGAGTTCTC 1343
QY 1855 ctctcacatgccagaggagagacag 1881
Db 1344 GTCGCACATAACGCGAGGGAGCACCG 1370
```

RESULT 4

```
US-08-465-590-5
; Sequence 5, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
;   Applicant: Georgopoulos, Katia A.
;   TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
;   NUMBER OF SEQUENCES: 164
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: LAHIVE & COCKFIELD
;   STREET: 60 STATE STREET, Suite 510
;   CITY: BOSTON
;   STATE: MASSACHUSETTS
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: AscII (text)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/465,590
;   FILING DATE: 05-JUN-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/238,212
;   FILING DATE: 02-MAY-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/121,438
;   FILING DATE: 14-SEP-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/946,233
;   FILING DATE: 14-SEP-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Myers, Paul L.
;   REGISTRATION NUMBER: 35,695
;   REFERENCE/DOCKET NUMBER: MPG-006C2DV
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)227-7400
;   TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2049 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 223..1776
;   US-08-465-590-5
```

```
Query Match 19.1%; Score 379.6; DB 3; Length 2049;
Best Local Similarity 61.8%; Pred. No. 2.1e-92;
Matches 771; Conservative 0; Mismatches 414; Indels 63; Gaps 8;
```

```
QY 700 cagcagacaaacccagtggaagatgaactgcagcgtgtgcgggttatcctgtaattagctt 759
Db 549 CATTCGACTCTCTTCAACGGAAACTAAAGTGTATATCTGTGGGATCGTTTGCATCGGGC 608
QY 760 caactgtgtggttcataagcgaagcattaccggcgaaccccggtccagtgtaaca 819
Db 609 CAAATGCTCATGTTTACAAAAGAGTCACTGTGTGAACGGCCTTTCCAGTGCACAA 668
QY 820 gtgcggggcatctttactcagaaggttaacctcctccctcatattaaactgcacacgg 879
Db 669 GTCTGGGGCTCTTTTACCCAGAAAGCACTCTCTCGGCACATCAAGCTGCACTCGG 728
QY 880 ggaataacaccttttaagtgtcacctctgcaactacgcatgccaaagagagatgcgtcac 939
Db 729 TGAGAAAGCCCTTCAAAATGCCATCTTTTGAACATATGCTCGCCGGGAGGACGCCCTCAC 788
QY 940 gggacaccttagacacattctgtgagagccgtacagtgatgagttctgcgaagaag 999
Db 789 CGGCCACCTTGAGGACCGACTCCGCTTGGTAAGCCTCAAAATGTGGATATTGTGGCCGGAG 848
QY 1000 ctacaagcagagaagctccctgtgagagcacaaggaacgctgcgagcttttttcagaa 1059
Db 849 CTATAAACAGCGAAGCTCTTTAGAGGAGCATAAAGAGCGATGCCAACACTACTTGGAAAG 908
QY 1060 cctcgacctgggggacg-----ctgcaagtggtggaggcaagacacatacaag----- 1107
Db 909 CATGGGCTTCCGGCGGTGTGCCCACTTAAAGGAAGAAACTTAACCAACAGAGATGCG 968
QY 1107 -----ccgaatgggaagttagagactctcctcctggacagattagcaagaa 1155
Db 969 AGAAGACCTGTGCAGATAGGAGCAGAGAGGTCCCTTGCTTGGAGACAGCTGCAAGCAA 1028
QY 1156 tgtggcgaagcaaaaaagctcgtcctcagaaattcatcgttgagaagcggcactgctt 1215
Db 1029 TGTGCGCCAAACGTAAGAGCTCTATGCTTCAGAAATTTCTTGGAGACAAAGTGCCTGTGCA 1088
```

Qy 1108 -----cgagatgggaagtagagagagctctcctcctggagacattagcaag 1152
Db 582 GCACGAAGACCTGTCAAGATAGGATCAGAGAGATCTCTCGTCTGGACAGACTAGCAAG 641
Qy 1153 caatgtgctaaagcaaaagctgctcctcagaaattcatcgttgagaagcgccactg 1212
Db 642 TAATGTCCCAACGCTAAGAGCTCTATGCTCAGAAATTTCTTGGGACAAAGGCGCTGTC 701
Qy 1213 ctctgatcccaactcaatcccggtacatgttacgagaaggaacagatgtagcagac 1272
Db 702 CGACAGCGCCTACGACAGTGC-----CACGTACGAGAAGGAGAACGAAATGATGAATC 755
Qy 1273 ccgagatgagcaagcaatcaataagcctacgtctatcagggctgaagccttcog 1332
Db 756 CCACGTGATGGACCAAGCCATCAACACGCCATCACTACCTTGGGGCGGAGTCCCTCGG 815
Qy 1333 ccccttagtccagatccgcctgctcccaaccttgagatgggtcccaagtacagcagtg 1392
Db 816 CCGCGTGTGACAGACGCCCGCG---GCGGTTCCGAGGTGGTCCCGGTCATCAGCCGAT 872
Qy 1393 gtaccacata-----gcacttactcgggcccagatatgcacatgggggcccccgca 1440
Db 873 GTACAGGTGTCAGAGCGCTCGAGGGCACCCCGCGCTCCAAACCACTCGGGCCAGGACAG 932
Qy 1441 ggagatgaagaagaacagatcctcctccagagagatcttgccttctgaacaggtct 1500
Db 933 CCGCGTGGAGTACCTGCTCTCTCCAAAGCCCAAGTTGGTGGCCCTCGGAGCGGAGGC 992
Qy 1501 gtcccccataacagtgcacagagctcacagacccagcagcaaccacagagatc--- 1557
Db 993 GTCCCGGAGCAACAGTCCCAAGACTCCAGGACACCCGAGACCAACACGAGGACGCG 1052
Qy 1557 ---gcaacatctctacagcaagaagcaagtggtctctccccagggcccccaatgggatgcc 1614
Db 1053 CAGCGGTCTTATCTACCTGACCAACACATCGCCGCGAGCGCGCAACGCTGTCG----- 1108
Qy 1615 tctctgaagaggtccctcgtcttttgaactctcaagccctcccatcgtcgtgag 1674
Db 1108 ----CTAAGGAGGAGACCGCGCTCTAGACCTGTGGCGCGCGCTCCGAGAACTCGCA 1163
Qy 1675 ggaactccatcaagtgtacaaagaagggaggtgagtggtgttctcgatgtgacca 1734
Db 1164 GGACGCGTCCCGTGGTGGTCCAGCACGAGCGGGGAGCAGATGAAGTGTACAGTGGCAACA 1223
Qy 1735 ctgcaacgtctctcctctagattatgtatgttccaccatcccaatgggggtgcatgttt 1794
Db 1224 CTGCGGGTGTCTTCTCTGGATCAGCTCATGTACACCATCCACATGGGCTGCCACGGCTT 1283
Qy 1795 cagtatcccttgatgttaacatgtgtgctatcgaagcaacgacgtcgtatgattctc 1854
Db 1284 CCGTGATCTTTTGTAGTGCAACATGTGGGCTTACCACAGCCAGGACCGGTACGAGTTCTC 1343
Qy 1855 ctctcacatcgccagagagacacag 1881
Db 1344 GTCCGACATAACCGGAGGAGCACCG 1370

RESULT 3

PCT-US95-09345-2
; Sequence 2, Application PC/TUS9509345
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS TRANSGENIC CELLS AND ANIMALS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/09345
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,300
; FILING DATE: 29-JULY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MGP-027PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1386 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1386
; PCT-US95-09345-2

Query Match 20.8%; Score 411.8; DB 5; Length 1386;
Best Local Similarity 63.3%; Pred. No. 4.2e-101;
Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;
Qy 700 cagcagaccacacagtgaggagatgaactgcgacgtgtgcgggttatcctgcattagctt 759
Db 162 CATTCGACTTCTTCAACGAAACTAAAGTGTGATATCTGTGGGATCATTTTCATCGGCC 221
Qy 760 caagctctgtgtgtcataagcgaagccataccggaagcgcgcgttccagtgtaatca 819
Db 222 CAATGTGCTCATGTGTTCAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCATCA 281
Qy 820 gtgcggggcatcttttactcagaaaggtaacctctcctcgtcatattaaactgcacacgg 879
Db 282 GTGCGGGGCTTCATTACCCAGAGGGCAACCTGTCTCGGCACATCAAGTGCATTCCGG 341
Qy 880 ggaataacaccttttaagtgtcaccctctgcaactagcagatgccaagaggagatgagctcac 939
Db 342 GGAGAAGCCCTTCAAAATGCCACCTCTGCAACTACGCTCGCCGCGGAGGAGCGCCCTCAC 401
Qy 940 gggacaccttaggacacattctgtggagaagccgtacaaagtgtgagttcttcggaagaag 999
Db 402 TGGCCACTTGAGGACGCACTCCGTGGTAAACCTTCACAAATGTGGATATTGTGGCCGAAG 461
Qy 1000 ctacaagcagagaagctccctggaggagcacaagaagcgtgccgagcgttttcttcagaa 1059
Db 462 CTATAAACAACGACGACGCTTTTAGAGGAACATAAAGAGGCGCTGCCACAACACTACTTGGAAAG 521
Qy 1060 ccttgacct---gggggacgtgcgaagtgtggaggaagacacacatcaaacg----- 1108
Db 522 CATGGGCGCTTCCGGGACACACTGTACCCAGTCACTAAAGAAGAAACTAAGCAGCAGTGAAT 581
Qy 1108 -----cgagatgggaagtagagagctcctcctcctggagacattagcaag 1152
Db 582 GCACGAAGACCTGTCAAGATAGGATCAGAGAGATCTCTCGTCTGGACAGACTAGCAAG 641

Qy 1060 ccctgacct--gggggacgtgcgaagtgtgaggaagacacatcaaaagc----- 1108
Db 744 CATGGCCCTTCGGGCACACGTGATCCCAAGTCAATTAAGAAAGAACTAAGCAGTGAAT 803
Qy 1108 -----cgaatggaagtgagagctctctctgacagattagcaag 1152
Db 804 GGCAGAACCTGTGCAAGATAGATCAGAGATCTCTGCTGGACAGACTAGCAAG 863
Qy 1153 caatgtgcttaagcgaagaaactgcgtcctcagaataatctcgtgagaagcgcactg 1212
Db 864 TAATGTGCCAAACGTAAAGCTATGCCCTCAGAAATTTCTGGGCAAGGCCCTGTGTC 923
Qy 1213 cttgagtgcaactacaatcccggtacatgtacgagaagagacagatgacagac 1272
Db 924 GCACAGCCCTTACGACAGTGC-----CACGTACGAGAAGGAGAACGAAATGATGAATC 977
Qy 1273 ccggtatgagcaagcaatcaataacgcatcagcttatctaggggtgaagccttcgc 1332
Db 978 CCAGTGATGACCAAGCCATCAACAGCCATCACTACCTGGGGCCGAGTCCCTGCG 1037
Qy 1333 cccttagtcagactccgctgtctccacctctgagatggtccagtcacagcagtgt 1392
Db 1038 CCGCTGTGTGACAGCGCCCGGG---CGGTCCGAGGTGTCTCCGGTCAATCAGCCGAT 1094
Qy 1393 gtaccccaata-----gcacttaactcgggcccagatgccaatgggggcccccgca 1440
Db 1095 GTACAGTGTGACAGCGCTCGGAGGACACCCCGCTCCACCACTCGGGCCAGGACAG 1154
Qy 1441 ggagatggaagaaacggatctctctcagagaagatctgtctctcagcagagctct 1500
Db 1155 CGCGTGTGAGTACCTGTGCTCTCCAAAGGCCAAGTGTGTCTCGGAGCGGAGGC 1214
Qy 1501 gtccccaataacagtgccaggaagctccacagacacagacagcaaccacagggatc--- 1557
Db 1215 GTCCCGGAGCAACAGCTGCCAAGACTCCACGACACCGAGAGCAACACGAGGAGCAGCG 1274
Qy 1557 --gcaacatctctaccagcaagaaacagctgtctctcccccagcccgccaatggatgcc 1614
Db 1275 CAGCGGTCTTATCTACCTGTACCAACCACTCGCCGCGAGCGGCAACGCGTGTGTCG--- 1330
Qy 1615 tcttgaagagaggtccctctctcttttgaactctctcgaagccctccctcctcctgag 1674
Db 1330 ---CTCAAGGAGGAGGACCGCGCTACGACCTCGTGGCGCGCCCTCCGAGAACTCGCA 1385
Qy 1675 ggaactcaataaagtatcaacaagaaggaggtgtgagatgtgttctgagatgaccca 1734
Db 1386 GGACGCGTTCGCGGTGTGTCAGCACCGCGGGGAGCAGATGAAGTGTACAAGTGCAGACA 1445
Qy 1735 ctgcccagctctctctctagattatgtgatttcaccatccacatggggtgcatgttt 1794
Db 1446 CTGCGGGTGTCTCTCTCTGATGACAGTGTATGATACCATCCACATGCGGCTGCCAGGCTT 1505
Qy 1795 ccgtgacccctttgagtgtaacatgtgtggtctatcgaagccacagctcgtatgattctc 1854
Db 1506 CGTGATCTTTTGTAGTGCACATGTGGGCTACCAAGCCAGGAGCCGGTACGAGTCTC 1565
Qy 1855 ctctcaatcgcagagagagagacagagccatgttgaagtgcac 1899
Db 1566 GTGCGACATACCGGAGGAGGACCGCTTCCACATGACGTGAAGC 1610

RESULT 2

US-08-465-590-3
; Sequence 3, Application US/08465590
; Patent No. 5824770
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510

CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,590
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,212
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,438
FILING DATE: 14-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,233
FILING DATE: 14-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,695
REFERENCE/DOCKET NUMBER: MPG-006C2DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1386
US-08-465-590-3

Query Match 20.8%; Score 411.8; DB 3; Length 1386;

Best Local Similarity 63.3%; Pred. No. 4.2e-101;

Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagaccacacagtggaagatgaactcgacgtgtgctgtggttattcctgcattagctt 759
Db 162 CATTCGACTTCTTAACGGAAACTAAAGTGTGATATCTGTGGATCATTTTCATCGGGCC 221
Qy 760 caacgtcttgatgttccataaaggaagccataccggaagccggaagcccggtccagtgtaatca 819
Db 222 CAATGTGCTCATGTGTTACAAAAGAACCCACACTGGAGAACGGCCCTTCCAGTGCAATCA 281
Qy 820 gtgcggggcactctttactcagaaggaagtaacctctcctcgtcatattaaactgcacacgg 879
Db 282 GTCGGGGCGCTCATTCACCCAGAGGCGCACTGCTCCGGCAGCATCAAGTGTGATTCGG 341
Qy 880 ggaataaaccttttaagtgtcacctctgcaactacgcatacccaagagagagatgcctcac 939
Db 342 GGAGAAGCCCTTCAATGTCCACCTCTGCAACTAGCGCTGCCCGGAGGAGCGCCCTCAC 401
Qy 940 ggacacacttagacacattctgtggaagacccgtacaaagtgtgagttctgcggaagaag 999
Db 402 TGCCACCTGTAGGACGCACTCCGTTGGTAAACCTCAAAATGTGGATATTGTGGCCGAAG 461
Qy 1000 ctacaagcagagaagctccctcgagagacacaggaagcagctgcgagcttttcttcagaa 1059
Db 462 CTATAACACGGAACGCTCTTTAGAGGAACATAAAGACCGCTGCCACAACTACTTGGAAAG 521
Qy 1060 ccttgacct--gggggacgctgcaagtgtgaggaagacacacatcaaaagc----- 1108
Db 522 CATGGCCCTTCGGGCACACACTGTGCCAGTCAATTAAGAAAGAACTAAGCAGTGAAT 581

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 06:55:19 ; Search time 44.29 Seconds
(without alignments)
4660.693 Million cell updates/sec

File:
Perfect score: 1984
Sequence: 1 cagagcgacacgcgctgg.....gaactcaaacccactcgag 1984

Scoring table: IDENTITY_NUC

Searched: 192659 seqs, 52021692 residues

..... : issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS9_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413.8	20.9	1611	5	PCT-US93-08743-3
2	411.8	20.8	1386	3	US-08-465-590-3
3	411.8	20.8	1386	3	US-08-465-590-3
4	379.6	19.1	2049	5	PCT-US95-09345-2
5	379.6	19.1	2049	5	US-08-465-590-5
6	346.4	17.5	1788	3	PCT-US95-09345-4
7	346.4	17.5	1788	3	US-08-465-590-2
8	346.4	17.5	1788	5	PCT-US93-08743-2
9	269.6	13.6	1004	3	PCT-US95-09345-1
10	269.6	13.6	1004	3	US-08-465-590-8
11	235.2	11.9	1170	3	PCT-US95-09345-7
12	235.2	11.9	1170	5	US-08-465-590-6
13	191.2	9.6	1296	3	PCT-US95-09345-5
14	191.2	9.6	1128	3	US-08-465-590-4
15	191.2	9.6	1296	5	PCT-US95-09345-3
16	191.2	9.6	1128	5	US-08-465-590-6
17	77.4	3.9	2031	4	PCT-US95-09345-6
18	68.2	3.4	7218	2	US-08-232-463-14
19	65.8	3.3	1892	4	US-08-933-750C-56
20	63	3.2	2133	3	US-08-820-170A-11
21	63	3.2	3754	3	US-08-820-170A-12
22	60.2	3.0	3720	1	US-08-074-967-1
23	60.2	3.0	3720	4	US-08-553-541B-1
24	60.2	3.0	3720	5	PCT-US94-06669-1
25	58.4	2.9	1309	4	US-08-933-750C-63
26	58.4	2.9	4252	5	PCT-US95-08429-4
27	53.6	2.7	3810	5	PCT-US95-08429-8
28	47.6	2.4	3132	3	US-08-224-482-3
29	46.4	2.3	2850	3	US-08-224-482-7
30	46.2	2.3	2811	3	US-08-040-548-31
31	46.2	2.3	2811	3	US-08-466-344-31
32	46.2	2.3	1440	3	US-08-224-482-5
33	44.4	2.2	267	3	US-08-040-548-14
34	44.4	2.2	3086	3	US-08-040-548-15
35	44.4	2.2	267	3	US-08-466-344-14
36	44.4	2.2	3086	3	US-08-466-344-15
37	44.4	2.2	3068	3	US-08-224-482-1

38	44	2.2	1161	4	US-08-616-857-1	Sequence 1, Appli
39	43.2	2.2	2043	4	US-08-398-590A-39	Sequence 39, Appli
40	43.2	2.2	5648	5	PCT-US96-03940-1	Sequence 1, Appli
41	43.2	2.2	756	5	PCT-US96-03940-2	Sequence 2, Appli
42	43.2	2.2	1407	5	PCT-US96-03940-3	Sequence 3, Appli
43	43.2	2.2	1090	5	PCT-US96-03940-4	Sequence 4, Appli
44	43.2	2.2	928	5	PCT-US96-03940-5	Sequence 5, Appli
45	43.2	2.2	1791	5	PCT-US96-03940-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
PCT-US93-08743-3
; Sequence 3, Application PC/TUS9308743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08743
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 14-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1611
PCT-US93-08743-3

Query Match		20.9%	Score 413.8;	DB 5;	Length 1611;
Best Local Similarity		63.18;	Pred. No. 1.3e-101;		
Matches 785;		Conservative 0;	Mismatches 397;	Indels 63;	Gaps 7;
Qy	700	cagcagacccaccagtggaagatgaactgcgacgtgtgcgggtttatctctcattagctt	759		
Db	384	CATTGCGACTTCTTAACGGAAACATAAGTGATATCTGTGGATCATTTGCGGGGCC	443		
Qy	760	caacgtcttgatggttcataagcgaagccatccacggcggaacgcgcgttccagtgtaaca	819		
Db	444	CAATGTGCTTATGTTTCAAAAAGAACCCACATGGAGAACGGCCCTTCCAGTGAATCA	503		
Qy	820	gtcgggggcactctttactcagaaggttaacctctcgtcatattaaactgcacacgg	879		
Db	504	GTGCGGGGCCCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCTTCGG	563		
Qy	880	ggaaaaaccttttaagtctacacctctgcaactagcatgccaagagagagatggcctcac	939		
Db	564	GGAGAAGCCCTTCAAAATGCCACCTGTGCACTACGCTGCGCGGAGGAGCGCCCTCAC	623		
Qy	940	gggacaccttgaggacacattctgtggagaagccgtacaagtgtgagttcttcggaagaag	999		
Db	624	TGGCCACCTGAGGAGCGCACTCCGTTGTTAAACCTCACAAAATGTGGATATTGTGGCCG	683		
Qy	1000	ctacaagcagagagctccctggaggagacacaaggaagcgtcgagctgtttttctcagaa	1059		
Db	684	CTATAACAGCGGACGCTTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGGAAG	743		

RESULT 15

T16065
ID T16065 standard; cDNA; 1004 BP.
AC T16065;
DT 09-MAY-1996 (first entry)
DE Ikaros cDNA.
KW Ikaros; transgene; transgenic animal; transgenic mouse;
KW immunocompromised; immune system disorder; nervous system disorder;
KW animal model; ss.
OS Not specified.
PN WO9604372-A1.
PD 15-FEB-1996.
PE 28-JUL-1995; U09345.
PR 29-JUL-1994; US-283300.
PA (GEMO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI: 96-129389/13.
DR P-PSDB: R92020.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
PS Disclosure: Page 71-73; 102pp; English.
CC A cDNA clone (T16065) coding for an Ikaros protein (R92020) has been identified. Transgenic animals, pref. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.
CC disorders.
SQ Sequence 1004 BP; 235 A; 328 C; 286 G; 155 T;

Query Match 13.6%; Score 269.6; DB 1; Length 1004;
Best Local Similarity 58.9%; Pred. No. 1.9e-58;
Matches 602; Conservative 0; Mismatches 369; Indels 51; Gaps 6;

QY 878 ggggaaaaacctttaagtgtcacctctgcaactacgcatcgaagagagagatgcctc 937
Db 1 GGAGAACGGCCCTCCAGTGGCAATCAGTGGCGGGCCCTCATTCACCCAGAGGGCAACTG 60
QY 938 acgggacaccttagacacattctgtggaagcgcgtacaaagtgtgattctgcggaaga 997
Db 61 CTCGGGCACATCAGCTGCTATCCGGGGAGAGCCCTTCAATGCCACCTCTGCAACTAC 120
QY 998 agctacaagcagagaagctccctgagagagcacaagaacgctgcgagcttttctcag 1057
Db 121 GCCTGCCCGCGGGAGCGCCCTCACTGGCCACCTGAGGACGCACCTCGTCAATAAGAA 180
QY 1058 aaccctgacctggggagcgtgcaagtgtgaggaagacacatcaaaagcagatggga 1117
Db 181 GAACATAAGCACAGTGAATGGC-----AGAAAGACCTGTGCAAGATAGGA 225
QY 1118 agtgagagagctctcgtctgacagattagcaagcaatgtggctaagcgaaaaagctcg 1177
Db 226 TCAGAGAGATCTCTGCTGTCGACAGACTAGCANGATATGTCGCCCAACAGTGAAGGCTCT 285
QY 1178 atgctcagaataatcattcggtgagaagcggcaactgtctcgatgccaactacaatccccggc 1237
Db 286 ATGGCTCAGAAATTTCTTGGGGACAAGGGCCCTGTCCGACACGCGCTACGACAGTGC--- 342
QY 1238 tacatgtacgagagaggaagacagatgatgagaccggagatgagaccaagccatcaat 1297
Db 342 --CAGTACGAGAGAGGAAGAAATGATGAAGTCCCACTGATGGACCAAGCCATCAAC 399
QY 1298 aacgcatcagctactaggggctgaagccttcgcgccttagtccagactccgcctgct 1357
Db 400 AACGCCATCACTACCTGGGGCCCGAGTCCCTGGCCCGCTGGTGCAGACGCCCCCGG-- 458
QY 1358 cccacctgtgagatggtcccgatcatcagcagtggtgtaccacct-----agca 1405
Db 458 -GCGGTTCCGAGGTGGTCCCGGTATCATCAGCCCGATGTACCACTGCACAGGCGCTCGGAG 516
QY 1406 cttactgggcccgatatgccaatggggggcccgagagatggaaaaaacgagatcctc 1465

Db 517 GGACACCCCGCGCTCCACACCACTCGGCCAGGACAGGCGCGTGAGTACCTGCTGCTC 576
QY 1466 ctgccagagaagatcttgccttgaacgaggtctgtcccccataaacagtgcaggac 1525
Db 577 TCCAAGGCCAAGTTGGTGGCCCTCGGAGGGGAGGCGTCCCGAGCAACAGCTGCCAAGAC 636
QY 1526 tcacagacacccagcaggaacccagggatc-----gccaacatctctaccagcaaac 1579
Db 637 TCCACGGACACCGAGAGCAACAACAGGAGCAGCGCGTCTTATCTACCTGACCAAC 696
QY 1580 cactgtgtctcccccagggccccaatggatgcctctctgaagaggtccctcgtct 1639
Db 697 CACATCGCCCGACGCGCGCAACGCGTGTG-----CTCAAGGAGGAGCACCAGCGCC 747
QY 1640 tttagaactcctcaagccccctccatctgctgagggaactccatcaaaagtacacaaa 1699
Db 748 TAGGACCTGTGCGCGCGCGCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCAC 807
QY 1700 gaaggaggagtgatggtgttctgatgtgaccaactgcacgtcctctctctagattat 1759
Db 808 AGCGGGGAGCAGATGAAGGTGTACAAAGTGCGAACACTGCCGGGTGCTCTTCTTGGATCAC 867
QY 1760 gtaatgttccacatccacatgggtgccatggttccgtgattcccttgaagtgaacatg 1819
Db 868 GTCATGTACACCATCCACATGGGCTGCCAGCGCTTCGGTATCCTTTTGTAGTCAACATG 927
QY 1820 tgtggctatcgaagccacgacgtcgtatgattctctctcacatcgccagagagagacac 1879
Db 928 TCGCGCTACCAACAGCCAGGACCGTACGAGTTCTCTGTCGCACATAACGCGAGGGAGCAC 987
QY 1880 ag 1881
Db 988 CG 989

Search completed: November 6, 1999, 02:38:40
Job time: 18371 sec

DB 1447 TATCAGCCAGGACAGGTACGAGTTCTCATCCATATACGCGGGGAGCATCGTTAC 1506

QY 1886 atgttgagtgagc 1899

DB 1507 CACCTGAGCTAAAC 1520

RESULT 14

V65968 ID V65968 standard; cDNA; 1788 bp.

AC V65968; 14-JAN-1999 (first entry)

DE Marine Ikaros encoding cDNA mik-2.

KW C3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;

KW differentiation marker; immune system; corpus striatum; AIDS;

KW Alzheimer's disease; ss.

OS Mus sp.

Key Location/Qualifiers

FT Key

CDS 223..1518

FI /*tag= a

II /product= "mik-2"

II /transl_except= (pos:385..387,aa:Gln)

PN US5824770-A.

PN 20-OCT-1998.

PN 05-JUN-1995; 465590.

PN 02-MAY-1994; US-238212.

PR 14-SEP-1992; US-946233.

PK 14-SEP-1993; US-121438.

PK 05-JUN-1995; US-465590.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K;

PI WPI: 98-582821/49.

DR P-PSDB; W72671.

PT Ikaros poly:peptide(s) - useful for treating disorders of immune

PT system or corpus striatum

PS Disclosure: Column 51-56; 11pp; English.

CC The present invention describes a purified peptide having at least one

CC of the following properties: (a) it stimulates transcription of a DNA

CC sequence under the control of a delta A element, an NFkB element or an

CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of

CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide

CC consensus sequence; (c) it competitively inhibits the binding of a

CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB

CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it

CC competitively inhibits Ikaros binding to Ikaros responsive elements; or

CC (e) it inhibits protein-protein interactions of transcriptional complexes

CC formed with naturally occurring Ikaros isoforms. The proteins, provided

CC that they stimulate gene transcription under the control of delta A

CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to

CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit binding of naturally occurring Ikaros isoforms to

CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or

CC inhibit protein-protein interactions of transcriptional complexes with

CC naturally occurring Ikaros isoforms, can be used to treat immune system

CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.

CC Alzheimer's disease. The present sequence encodes a specifically

CC claimed mouse Ikaros protein.

SQ Sequence 1788 BP; 437 A; 486 C; 495 G; 370 T;

Query Match 17.58; Score 346.4; DB 1; Length 1788;

Best Local Similarity 62.08; Pred. No. 1.4e-77;

Matches 715; Conservative 0; Mismatches 376; Indels 63; Gaps 8;

QY 794 ggcgaacgccgttccagtgtaacagtcgagggcattttactcagaagtaaacctc 853

DB 382 GGTGAACGGCTTTCCAGTGCACAGCTCTGGGCGCTTCTTACCCAGAAAGGCAACCTC 441

QY 854 ctccgtcatattaaactgcacacgggggaaacaccttttaagtgtcacctctgcaactac 913

DB 442 CTGCGGCACATCAGCTGCACTCGGGTGAGAGCCCTTCAATGCCATCTTGCACATAT 501

QY 914 gcatgccaaagagagatcgctcagcgagacaccttaggacacattctgtgagagccg 973

DB 502 GCGTCCGCGGAGGAGCCCTCTACCGGCCACCTGAGGACGACTCCCTTGGTAAGCCT 561

QY 974 tacaagtgtgagttctgcggaagaagactacaagcagaagctccctggagagacacaag 1033

DB 562 CACAAATGTGGATATTGTGCGCGAGCTATAAACAGCGAAGCTCTTTAGAGAGGATAAA 621

QY 1034 gaacgctcgcgagcttttctcagaacctgacctgggggagcgtctcgaagtgtg 1087

DB 622 GAGCGATGCCACAACACTACTTGGAAACCATGGGCTTCCGGCGGTGCGCCAGTCATTAG 681

QY 1088 gaggcgaagacacatcaaaagctcgaagtgggaagtggagagagct 1129

DB 682 GAAGAACTAAACCAACACAGAGATGCGAGAAGACCTGTGCAAGATAGGACGAGAGGTCC 741

QY 1130 ctccgtctgcagagattagcaagcaatgtgtaagcgaagaaagctcgatgctcagaaa 1189

DB 742 CTGTCTCTGGACAGGCTGGCAAGCAATGCGCCAAACGTAAGAGCTCTATGCTCAGAAA 801

QY 1190 ttcatcggtgagaagcggcactgcttccgacacacacacacacacacacacacacacac 1249

DB 802 TTCTTTGGAGACAGTGCCTGTGACACATGCCCTATGACAGTCCCAACTATGAGAAGGAG 861

QY 1250 aaggagaacagatgatgcagaccggatgtagccaccagccatcaataacgccatcagc 1309

DB 862 -----GATATGATGACATCCACGCTGATGGACAGGCCATCAACAATGCCATCAAC 912

QY 1310 tatctagggctgaagccttcgccccttagtccagactccgctgctccaccctctgag 1369

DB 913 TACCTGGGGCTGAGTCCCTGCGCCATTTGGTGACAGACACCCCGGT---AGCTCCGAG 969

QY 1370 atggtcccatcatcagcagtggtaccaccatagcaacttactcgggcccagatgccaatg 1429

DB 970 GTGGTGCCAGTCATCAGTCCATGTACAGCTGCACAAAGCCCTTCAGATGGCCCCCA 1029

QY 1430 ggg-----gcccgagagagatggaaaagaaacgagctcctcctgccagagag 1477

DB 1030 CGGTCCCAACCAATTCACAGCAGGACGCGGTGATTAACCTTGTCTGCTCCCAAGGCCAAG 1089

QY 1478 atcttgctcttgaaagaggtctgtcccccataaacagtgcccagagctccacacaccc 1537

DB 1090 TCTGTCTCATGGAGGAGAGGCGCTCCCGAGCAACAGCTGCCAAGACTCCACAGATACA 1149

QY 1538 gacagcaacacagag---atcgccaacatctctaccagcaaacgacggtggtcctcccc 1594

DB 1150 GAGAGCAACGCGAGGACAGCGAGCGGCTTATCTACTTAACCAACCACTCAACCCG 1209

QY 1595 caggccgcaatggagtgctcttctgaaggaggtccctcgtctcttttgaactcctcaag 1654

DB 1210 CATGCACGAATGGGTGCTCTC---AAGGAGGAGCAGCGGCGCTTACGAGGTGCTGAGG 1266

QY 1655 cccctcccatctgctcagagagctccatcaagtgatcaacaagaaggaggtgatg 1714

DB 1267 GCGGCTCTAGAGAACTCGCAGGATGCCCTTCCGTGTGGTCAGCAGAGTGCGGAGCAGCTG 1326

QY 1715 gatgtgtttcctgatgaccactgcccacgtccctctctctctctctctctctcaccatc 1774

DB 1327 AAGGTGTACAAGTGCAGAACACTGCGCGGTCTCTCTCTGGATCACGTCATGATACCAT 1386

QY 1775 caccatg-----gggtgccatggttcccgatgacccctttgagtgtaacatggtgagc 1825

DB 1387 CACATGGGTGCCATGGCTGCCATGGCTTTCGGGATCCCTTTGAGTGTAAACATGTGTGT 1446

QY 1826 tatcgagccagatcgatgattctcctctcctcctcctcctcctcctcctcctcctcctc 1885

DB 1447 TATCAGCCAGGACAGTACGAGTTCTCATCCATATCACCGGGGGAGCATCTGTTAC 1506

QY 1886 atgttgagtgagc 1899

DB 1507 CACCTGAGCTAAAC 1520

QY 1538 gacagcaaccacagagg---atcgccaacatctctaccagcaaaagccacagctggtctccccc 1594
DB 1150 GAGAGCAACGCGGAGGAAGCAGCGCAGCGGCTTATCTACTAACCAACACATCAACCCG 1209
QY 1595 caggcccgcaatgggagtgctcttctgaagggaggtccctcgtctcttttgaacctctcaag 1654
DB 1210 CATGACGCAATGGCTGCTCTC---AAGGAGGACGACGCGCTTACGAGGTGTGAGG 1266
QY 1655 cccctcccatctgctgagggaactccatcaaaagtatcaacaaagaagggaggtgatg 1714
DB 1267 GCGGCTCAGAGAACTCGCAGGATGCTTCCGTGTGGTCAGCACGAGTGGCAGCAGCTG 1326
QY 1715 gatgtttcgtatgacacatgccaatgccaatgctctctctctctagattatgtgttccaccatc 1774
DB 1327 AAGGTGTACAAGTGGCAACACTGCCGCGTCTTCTCTGGATCAGCTCATGTATACCAT 1386
QY 1715 cacatg-----gggtgccatggtttccgttgatccctttgagtgtaacatggtggtgc 1825
DB 1387 CACATGGGCTGCCATGGCTGCCATGGCTTTCGGATCCCTTTGAGTGTAAACATGTGTGT 1446
QY 1826 tatcgaagccacgatcgtatgattctctctctcacaatcgcagagagagacagagacc 1885
DB 1447 TATCAGACGACGACAGGTACGAGTCTCATCCATATCACGCGGGGGAGCATCGTTAC 1506
QY 1886 atgttgaagtgaac 1899
DB 1507 CACCTGAGCTAAAC 1520

RESULT 13

T16059
ID T16059 standard; CDNA; 1788 BP.
AC T16059;
DT 08-MAY-1996 (first entry)
DE Murine Ikaros CDNA mik-2.
KW Ikaros; transgene; transgenic animal; transgenic mouse; mik-2;
KW immunocomprised; immune system disorder; nervous system disorder;
KW animal model; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 233..1518
FT /*tag= a
FT /transl_except= pos:385..388;_aa:Gln
PN W09604372-N1.
PD 15-FEB-1996.
PF 28-JUL-1995; U09345.
PR 29-JUL-1994; US-283300.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI: 96-129385/13.
DR P-PSDB; R92014.
PT Transgenic rodent having Ikaros trans-gene (pref. mutated) - is severely immuno-compromised and can be used as model to determine effects of treatment for immune and nervous system disorders
PS Disclosure: Fig 1: 102pp; English.
CC CDNA clones (T16059 and T16061-64) encode different isoforms, mik-2, mik-3, mik-1, mik-4 and mik-5 (R92014 and R92016-19, respectively), of the mouse Ikaros protein, a master regulator of hematopoietic differentiation. They were isolated from a mature T-cell line E-14 CDNA library. The isoforms arise by differential splicing of the Ikaros genomic locus. All include exons E1/2 and E7, but have different combinations of exons E3-E6 encoding a zinc finger domain.
CC Transgenic animals, pref. mice, having a mutated Ikaros transgene, esp. a mutation that alters the DNA binding domain of the Ikaros protein, are used as models to determine the effects of treatments for immune or nervous system disorders.
SQ Sequence 1788 BP; 437 A; 486 C; 495 G; 370 T;

Query Match 17.5%; Score 346.4; DB 1; Length 1788;
Best Local Similarity 62.0%; Pred. No. 1.4e-77;
Matches 715; Conservative 0; Mismatches 376; Indels 63; Gaps 8;

QY 794 ggcgaacgcccgttcccaagttaatacagtcggggcatcttttactcagaagaaggtaaacctc 853
DB 382 GGTGAACGGCTTTCCAGTGAACAGCTGTGGGCCCTCTTTACCCAGAAAGGCAACCTC 441
QY 854 ctccgtcatataataacatgcacacgagggaacacacacacacacacacacacacacacacac 913
DB 442 CTGGGCAACATCAAGCTCACTCGGTGAGAGGCCCTTCAATGCCATCTTTTGAACATAT 501
QY 914 gcatgccaaaagagagatgcgtcacgggagcacaccttagcacacatcttctgtggagagcgc 973
DB 502 GCTTGCCTCGGAGGAGACGCCCTCACGGCCACCTGAGAGCGCACTCCGTTGGTAAGCCT 561
QY 974 tacaagtgtgagttcttcggaagaagctacaaagcagaagaagctccctctggagagacacaag 1033
DB 562 CACAAATGTGGATATTGTGGCCGAGGCTATAAACAGCGAAGCTCTTTAGAGAGCATATAA 621
QY 1034 gaacgtcccgagcttttcttcagaacacctgacctgggggagc-----ctgcaagtgtg 1087
DB 622 GAGCATGCCACACTACTTGGAAAGCATGGCCCTTCCGGCGGTGTCGCCAGTCATTAAAG 681
QY 1088 gagcgaagacacatcaaaag-----ccgagatggggaagtgaagagct 1129
DB 682 GAAGAACTAACCAACAGAGATGCGCAGACACCTGTCAAGATAGGAGCAGAGAGTCC 741
QY 1130 ctgctctggacagattagcaagaatgtggctaaagcgaagaagctccctctggagagacacaag 1189
DB 742 CTTGTCTTGGACAGGCTGGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCCCTCAGAAA 801
QY 1190 ttcatcgtgtgagaagcggcactgcttgcgtcccaactacatacctccggctacatatcagag 1249
DB 802 TTCTTGGAGACAGTGCCTGTACAGATGCCCTATGACATGCCCACTATGAGAAGGAG 861
QY 1250 aaggagaacgagatgatgacacccggatgatgacaaagccatcaataaaccgcatcagc 1309
DB 862 -----GATATGATGACATCCACGCTGATGACACAGGCCATCAACAATGCCATCAAC 912
QY 1310 tatctaggggtgaagccttccgccccttagtccagactccgctctccacacctgag 1369
DB 913 TACCTGGGGGTGAGTCCCTCGCCCATTTGTCGACACACCCCGGT---AGTCCGAG 969
QY 1370 atggtcccgatcagcagtgatccccaatagcacttactcggccgctatcccaatg 1429
DB 970 GTGTGCCAGTCATCAGTCCATGTACCACTGTCAGCTGCACAAAGCCCTCTAGATGGCCCCCA 1029
QY 1430 ggg-----gccccgcaggagatgaaagaacggatcctcctgccagagaga 1477
DB 1030 CGGTCCAAACCATTCAGCACAGACGCGCTGGATTAACTTGCTGCTGTCTCCAAGGCCAAG 1089
QY 1478 atcttgccttctgaacaggtctgttcccccaataaacagtgccccaggactccacagacacc 1537
DB 1090 TCTGTGTATCGGAGCGAGAGCGCTTCCCGAGCAACAGCTGCCAAGACTCCACAGATACA 1149
QY 1538 gacagcaaccacaggg---atcgccaacatctctaccagcaaaagccacgctggtctccccc 1594
DB 1150 GAGACACACGCGGAGGAACACGCGCGCTTATCTACTAACCAACCAACATCAACCCG 1209
QY 1595 caggcccgcaatgggagtgctcttctgaagggaggtccctcgtctcttttgaacctctcaag 1654
DB 1210 CATGACGCAATGGCTGCTCTC---AAGGAGGAGCAGCGCGCTTACGAGGTGTGAGG 1266
QY 1655 cccctcccatctgctgagggaactccatcaaaagtatcaacaaagaagggaggtgatg 1714
DB 1267 GCGGCTCAGAGAACTCGCAGGATGCTTCCGTGTGGTCAGCAGAGGTGGCGGAGCAGCTG 1326
QY 1715 gatgtttcgtatgacacatgccaatgccaatgctctctctctagattatgtgttccaccatc 1774
DB 1327 AAGGTGTACAAGTGGCAACACTGCCGCGTCTCTTCTGGATCAGCTCATGTATACCAT 1386
QY 1775 cacatg-----gggtgccatggtttccgttgatccctttgagtgtaacatggtggtgc 1825
DB 1387 CACATGGGCTGCCATGGCTGCCATGGCTTTCGGGATCCCTTTGAGTGTAAACATGTGTGT 1446
QY 1826 tategaagccacgatcgtatgattctctctctcacaatcgcagagagagagacagagcc 1885

Db 742 CTTGCTCTGGCAGGCTGGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCTCTCAGAAA 801
 Qy 1190 ttcatcggtgagaagcgactgtcttcgatgccaactacaaatccccggctacatgtacag 1249
 Db 802 TTTCTTGAGACAAGTGCCTGTGACACATGCCCTATGACAGTGCCTACACTATGAGAGGAG 861
 Qy 1250 aaggagaacagatgatgcagacccggatgatggaccaaacccatcaataaagccatcagc 1309
 Db 862 -----GATATGATGACATCCACGCTGATGACACGCCCTCAACATGCCATCAAC 912
 Qy 1310 tatcagggctgaagccttcgccccttagctccagactccgctgctccacctctgag 1369
 Db 913 TACTTGGGGCTGAGTCCCTGCGCCCATTTGGTCAGACACCCCCCGT---AGCTCCGAG 969
 Qy 1370 atggtccagctcatcagcagtggtgtaccoccatagcacttactcggggccgatgccaatg 1429
 Db 970 GTGGTGCACATCATCAGCTCCATGTACCATGCTGCACAAAGCCGCCCTCAGATGGCCCCCA 1029
 Qy 1430 ggg-----gcccgcagagatggaaagaaacgagatcctctcgcagagag 1477
 Db 1030 CGGTCCAAACCATTCAGCAGGACGCGGTGGATAACTTGTCTGCTGTCTCCAAAGGCCAAG 1089
 Qy 1478 atcttgcttctgaacgaggtctgtcccccaataacagtgccacgaggtccacagacacc 1537
 Db 1090 TCTGTGTCTCGGAGGAGGAGGCGCTCCCGGAGACAGCTGCCAGACTGCCACAGATACA 1149
 Qy 1538 gacagcaaccacaggg---atcgcccaacatctctaccagcaaaagccacggtggtccctccc 1594
 Db 1150 GAGAGCAACCGGAGGAAACAGCGCAGCGCTTATCTACCTAACCAACCAACATCAACCCG 1209
 Qy 1595 caggccgcgaatggatgctcttctgaagaggctccctgcctcttctgaactcctcaag 1654
 Db 1210 CATGCAGCAATGGCTGGCTCTC---AAGGAGGAGCAGCGCCCTACGAGGTGCTGAGG 1266
 Qy 1655 cccctccctctgctgagggactcctcaaaagtgtatcaacaagaaggggaggtgatg 1714
 Db 1267 GCGGCTCAGAGAACTCGCAGGATGCTTCGGTGTGTCAGCAGGAGTGCGGAGAGCTG 1326
 Qy 1715 gatgtttctgatgtgaccactgccacgtctcttctctcttcttagatgtatgttccaccatc 1774
 Db 1327 AAGGTGTACAAGTGCAGAACTGCGCGGTGCTCTTCTGATGATCAGCTCATGTATACATT 1386
 Qy 1775 cacatg-----gggtccatggttccggtgatccctttgagtgtaacatgtgtgc 1825
 Db 1387 CACATGGGCTGCCATGCTGCCATGGCTTTCGGGATCCCTTTGAGTGTAAATGTGTGT 1446
 Qy 1826 tatcgaagccacgatcgtatgattctctctcacatcccgagagagagagacagagcc 1885
 Db 1447 TATCACAGCCAGGACAGGTACGAGTCTCATCCATATACCCGGGGGAGCATCGTTAC 1506
 Qy 1886 atgttgaagtgaac 1899
 Db 1507 CACCTGAGCTAAAC 1520

RESULT 12

y44579
 ID Q44979 standard; cDNA; 1788 BP.
 AC Q44979;
 NC 21-OCT-1994 (first entry)
 NT wine Ikaros gene.
 KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
 KW corpus striatum; regulatory gene; ss.
 OS Mus musculus.
 FH key Location/Qualifiers
 FT 223..1518
 FT cds /*tag= a
 FT /product= Ikaros protein.
 FT 385..387
 FT misc_feature /*tag= b
 FT /transl_except= GAA encodes Glutamine.
 PN WO9406814-A.
 PD 31-MAR-1994.

PF 14-SEP-1993; U08743.
 PR 14-SEP-1992; US-946233.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI: 94-118387/14.
 DR P-PSDB; R46963.
 DR T-cell pathway regulatory gene, Ikaros - encodes family of unique
 PT zinc finger proteins, useful for treating immune system disorders
 PS Claim 13; Page 41-43; 112pp; English.
 CC The Ikaros gene encodes a zinc finger protein which can be used in a
 CC therapeutic composition to treat animals with an immune system
 CC disorder. It may also be used for assessing whether a subject is at
 CC risk for an immune disorder. It is of particular use in treating a
 CC disorder of the corpus striatum.
 SQ Sequence 1788 BP; 437 A; 486 C; 495 G; 370 T;

Query Match 17.5%; Score 346.4; DB 1; Length 1788;
 Best Local Similarity 62.0%; Pred. No. 1.4e-77;
 Matches 715; Conservative 0; Mismatches 376; Indels 63; Gaps 8;

Qy 794 ggcgaacgcccgttccagtgtaatcagtcggggcatcttttactcagaaggttaacctc 853
 Db 382 GTGTAACGCCCTTCCAGTGCACACAGCTGCGGCTCTTTACCCAGAAAGCCACCTC 441
 Qy 854 ctccgtcatattaaactgcacacgggggaaacaccttttaagtgtcacctctgcaactac 913
 Db 442 CTGCGGCACATCAAGTGCACCTCGGTGAGAAGCCCTTCAAAATGCCATCTTTGCAACTAT 501
 Qy 914 gcatgcaaaagagagatgcgtctcaggggacaccttagacacattctgtggagaagcgg 973
 Db 502 GCGTCCGCGCGGAGGAGCGCCCTCACCGGCCACCTGAGGACGCACCTCCGTTGGTAAGCCT 561
 Qy 974 tacaagtgtgactctcgggaagactcaagcagagaagctccctcgtgagagacacaag 1033
 Db 562 CACAATGTGGATATGTGCGCGGAGCTATAAACAGCAGAGCTCTTTAGAGAGGATATAA 621
 Qy 1034 gaacgtgcccagcttttcttcagaacctgacctgggggagc-----ctgcaagtgtg 1087
 Db 622 GAGCGATGCCACAACATCTTGGAAAGCATGGGCTTCCGGCGGTGTGCCAGCTATTAAAG 681
 Qy 1088 gaggcaagacacataaag-----ccgagatgggaagttagagagct 1129
 Db 682 GAAGAAACTTAACCAACAGAGATGCGAAGACCTGTGCAAGATAGGACGACAGAGGTCC 741
 Qy 1130 ctctcctctgacagattagcaagcaatgtgctaagcgaaagctcgtgagcctcagaaa 1189
 Db 742 CTGTCTCTGGACAGGCTGCGCAGCAATGTGCGCAACAGTAAAGAGCTCTATGCTCAGAAA 801
 Qy 1190 ttcacgtgtgagaagcggcactgcttcgatgccaactacaatccccgggtctacatgtacgag 1249
 Db 802 TTTCTTGGAGACAAGTGCCTGTGACACATGCCCTATGACAGTGCCAACTATGAGAAGGAG 861
 Qy 1250 aagggaacagagatgcagacccggatgatggaccaagccatcaataaacccatcagc 1309
 Db 862 -----GATATGATGACATCCACGCTGATGACACGCCCTCAACATGCCATCAAC 912
 Qy 1310 tatctaggggctgaagccttcgccccttagtccagactccgctgctccccctctgag 1369
 Db 913 TACTTGGGGCTGAGTCCCTGCGCCCATTTGGTGTGAGACACCCCCCGT---AGCTCCGAG 969
 Qy 1370 atggtccagctcatcagcagtggtgtaccoccatagcacttactcggggccgatgccaatg 1429
 Db 970 GTGGTGCACATCATCAGCTCCATGTACCATGCTGCACAAAGCCGCCCTCAGATGGCCCCCA 1029
 Qy 1430 ggg-----gcccgcagagatggaaagaaacgagatcctctcgcagagag 1477
 Db 1030 CGGTCCAAACCATTCAGCAGGACGCGGTGGATAACTTGTCTGCTGTCTCCAAAGGCCAAG 1089
 Qy 1478 atcttgcttctgaacgaggtctgtcccccaataacagtgccacgaggtccacagacacc 1537
 Db 1090 TCTGTGTCTCGGAGGAGGAGGCGCTCCCGGAGACAGCTGCCAGACTGCCACAGATACA 1149

```

Db 969 AGAAGACCTGTCAGATAGGAGCAGAGAGGTCCTTGTCTGGACAGCTGGCAAGCAA 1028
Qy 1156 tggctgaagcgaagaaagctgctcctcagaataatcctggtgagaagcgcaactgtt 1215
Db 1029 TGTGCCAAACGTAAGAGCTCTATGCTCAGAAATTTCTTGAGACAAGTGCCTGTGAGA 1088
Qy 1216 cgaatccaaactacaatcccggtacatgtacgagaagagaacgagatgatgcagaccg 1275
Db 1089 CATGCCCTATGACGTGCAACTATG-----AGAAGGAGGATATGATGACATCCCA 1139
Qy 1276 gatgatgacaaagccatcaataacgcgcacagctataggggtgagccttcgcgcc 1335
Db 1140 CGTGATGACCAAGGCGCAATCAAAATGCCATCAACTACCTGGGGGTGAGTCCCTCGGCC 1199
Qy 1336 cttagtcagactccgctgctccacactctgagatggtcccaagtcacagcagtgta 1395
Db 1200 ATTGGTGACAGACCCCGGT---AGTCCGAGGTGGTGGCAGTCATCAGTCCATGTA 1256
Qy 1396 ccccatagcacttactcgggcccgaatatgccaatggggccc-----ccgcagga 1443
Db 1257 CCAGCTGCAAGACCCCGCTCAGATGGCCCGCCACGGTCCAAACCATTCAGACAGGAGCG 1316
Qy 1444 gatgaaagaagacgactcctcctccagagagaagatcttgccttgacgaggtctgc 1503
Db 1317 CGTGATTAATGCTGCTGCTGCTCAAGGCCAAAGTCTGTGTATCGGAGCGAGGCGCTC 1376
Qy 1504 ccccaataacagtcgaggaagctccacagacacgacagcaaccagagg---atcgcca 1560
Db 1377 CCCGAGCAACAGCTGCCAAGACTCCACAGATACAGAGCAACGCGGAGGACAGCGAG 1436
Qy 1561 acatctaccgcaaaagccagtcggtgctcccccagggcccgcaatggatgctcttct 1620
Db 1437 CGGCTTATCTACCTAAGCAACACATCAACCCGCATGCAGCAATGGCTGGC---TCT 1493
Qy 1621 gaagaggtcctcgtctcttttgactcctcgaagccctccatctcctcgtgaggactc 1680
Db 1494 CAAGGAGGAGCAGCGCGCTACGAGGTGCTGAGGGCGGCTCAGAGAACTCGCAGAGTGC 1553
Qy 1681 catcaagtgtacaagaagggaggtgatgatgtgttccatgtgacacactgcca 1740
Db 1554 CTTCGCTGTGTCAGCAGAGTGGCGAGCAGCTGAAGTGTACAGTGGACACTGGCG 1613
Qy 1741 gctcctctcctagatattatgtatgttccacatccacatg-----gggtgcatgg 1791
Db 1614 CGTGCTCTCTGATCAGCTCATGTATACCATTCACATGGGCTGCCATGGCTGCCATGG 1673
Qy 1792 ttccgtgatcccttgatgtaacatgtgtgctatgagcagacgatcgctatgatt 1851
Db 1674 CTTTCGGGATCCCTTTGAGTGTAAATGTGTGTGTTATCAGACCCAGGACAGTACGAGTT 1733
Qy 1852 ctctctcacatcgccagaggagacagagccatgttgagagtgc 1899
Db 1734 CTCATCCCATATCACGCGGGGGAGCATCGTTACCACTGAGCTAAAC 1781

RESULT 11
v42805
AC v42805 standard; cDNA; 1788 BP.
DE 11-JAN-1999 (first entry)
KW Mouse Ikaro isoform mik-2 cDNA.
KW Ikaro; mik-2; transcription factor; mouse; lymphocyte;
KW cell differentiation; T cell; cancer; immunodeficiency;
KW Alzheimer's disease; therapy; diagnosis; ss.
OS Mus sp.
FH Key
FT Location/Qualifiers
FT 223..3518
FT /*tag= a
FT /transl_except= (pos:385,,387, aa:Gln)
PN CA2194256-A.
PD 05-MAR-1998.
PF 02-JAN-1997; 194256.

```

```

PR 05-SEP-1996; US-711417.
PA (GEHO ) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI; 98-378292/33.
DR P-PSDB; W70963.
PT New nucleic acid encoding Ikaro protein involved in early
PT differentiation of lymphocytes - existing in several isoforms, and
PT related products, used to treat e.g. immune diseases or cancer and
PT to control cell differentiation
PS Claim 7; Page 68-70; 158pp; English.
CC This is the nucleotide of mouse Ikaro cDNA (isoform mik-2) that
CC codes for a 431-amino acid zinc finger protein (see W70963) that is
CC involved in the early differentiation of lymphocytes. A cDNA
CC library constructed from the mature murine T cell line E14 was
CC screened with a multimerised oligonucleotide (see V42829) derived
CC from a protein binding site (see V42804) of the CD3-delta enhancer
CC to identify T cell specific sequences that bind and mediate
CC enhancer function. An isolated clone was designated Ikaro and
CC contained the 1788 bp sequence. A 300 bp 3'-terminal segment of
CC this was used to identify 4 other Ikaro isoforms (see V42807-10).
CC Different isoforms arise by differential splicing of Ikaro gene
CC transcripts. Isoform mik-2 lacks exon 3. It is expressed in
CC embryonic liver, thymus and brain, but only in thymus and spleen
CC after birth. The Ikaro gene is located at the proximal arm of
CC murine chromosome 11. Ikaro proteins are suggested to play a role
CC as a genetic switch regulating entry into the T cell lineage. The
CC murine and human (see V42806, V42811 and V42840) Ikaro sequences
CC are very similar. The invention provides Ikaro nucleic acids,
CC vectors and host cells expressing Ikaro proteins. These are used
CC to treat T and B cell diseases (e.g. immune deficiencies caused by
CC drugs, radiation or cancers), to control expression of heterologous
CC genes placed under control of an Ikaro-responsive element, to
CC treat nervous system diseases (e.g. Alzheimer's disease) and to
CC modulate cell division, amplification or differentiation, especially
CC in haematopoietic cells. Some Ikaro isoforms are antagonistic of
CC others and may be used to inhibit interaction with DNA sequences.
CC The same effect can be achieved with Ikaro-binding oligonucleotides.
CC Examining the expression of the Ikaro gene, or its allelic
CC structure, can be used to assess risk of acquiring the above
CC diseases.
SQ Sequence 1788 BP; 437 A; 488 C; 493 G; 370 T;

```

```

Query Match 17.5%; Score 348; DB 1; Length 1788;
Best Local Similarity 62.0%; Pred. No. 5.5e-78;
Matches 716; Conservative 0; Mismatches 375; Indels 63; Gaps 8;

Qy 794 ggcgaagcccccctccagtgtaatacagtcgaggcgatcttttactcagaaggaacctc 853
Db 382 GGTGAACGGGCTTTCCAGTGCACACAGCTCTGGGGCTCTTTTACCAGAAAGGCAACCTC 441
Qy 854 ctccgtcatataaactgcacacgggggaaaaacaccttttaagtctcacctctgcaactac 913
Db 442 CTGGCGGCATCAAGCTGCACCTCGGTGAGNAGCCCTTCAATGCCATCTTTGCACTAT 501
Qy 914 gcatgccaaaggagatgcgctcacgggacacaccttaggacacattctgtggagaagccg 973
Db 502 GCCTGCCCGCGGAGGAGCGCCCTCACCGGCACCTGAGGAGCGCACTCCGTGGTAAAGCCT 561
Qy 974 tacaagtgtgagttctcggaagaagctacaagcagagaagctccctcgaggagacacaag 1033
Db 562 CACAAATGTGGATATTGTGCCGGAGCTATAAACAGGCAAGCTCTTTAGAGAGCATAAA 621
Qy 1034 gaacgctgcgagcttttcttcagaacccctgacctgggggagc-----ctcgaagtgtg 1087
Db 522 GAGGATGCCCAACTACTTGGAAAGCATGGGCCCTTCGGGGCGTGTGCCCATGCTATTAG 681
Qy 1088 gaggcaagacacatcaaaag-----ccgagatgggaagtgaagagct 1129
Db 682 GAAGAAACTAACCAACAGAGATGGCAGAGACCTGTGTAAGATAGGAGCAGAGAGTCC 741
Qy 1130 ctgctctggacagattagcaagcaatgtggttaagcgaaaaagctcgatgcctcagaaa 1199

```


Db 993 GTCCCCGAGCAACAGCTGCCAACACTCCACGGACCGGACCGAGAGCAACACGAGGAGCAGCG 1052
 Qy 1557 --gcaacatctctaccagcaagccacgtggtctctccccagccgcgaatgggagtc 1614
 Db 1053 CAGCGGTCTTATCTAGCTACCAACACATCGCCGCGCGCAACGCGGTGTGCG----- 1108
 Qy 1615 tcttcgaaggaggtccctcgctctttttaaactctcaagccctcccatctgctgag 1674
 Db 1108 ----CTCAGGAGGAGACCGCGCTTACGACCTGCTGCGCGCGCTCCGAGAACTCGCA 1163
 Qy 1675 ggactccataaagtatcaacaagaaggaggtgagtggttttctcagatgacga 1734
 Db 1164 GGACGGCTCCGCGTGTGTCAGCACCAGCGGGGAGCAGATGAAGTGTACAAGTGGCA 1223
 Qy 1735 ctgcaagctctctctagattatgtatgttcaaccatccacatgggtgcaatggttt 1794
 Db 1224 CTGCGGGGTCTCTCTGATCAGCTCATGTACACATCCATGGGTGCGCACGGCTT 1283
 Qy 1795 ccgtgaccccttggatgaacatgtgtggtctatgaagccagatcgctatgattctc 1854
 Db 1284 CCGTATCTTTTGTGTCACATGTGCGGTACACAGCCAGGACCGGTACGAGTCTC 1343
 Qy 1855 ctctcatcgccagaggagagcacag 1881
 Db 1344 GTGCGACATAACGCGAGGGGAGCACCG 1370

RESULT 6

ID V42806
 AC V42806;

DT 11-JAN-1999 (first entry)

DE Human Ikaros isoform hix-1 cDNA.

KW Ikaros; hix-1; transcription factor; human; lymphocyte;

KW cell differentiation; T cell; cancer; immunodeficiency;

KW Alzheimer's disease; therapy; diagnosis; ss.

OS Homo sapiens.

FH Key

FT Location/Qualifiers

FT exon 1..255

FT /*tag= a

FT /number= Ex3

FT 256..423

FT /*tag= b

FT /number= Ex4

FT 424..549

FT /*tag= c

FT /number= Ex5

FT 550..684

FT /*tag= d

FT /number= Ex6

FT 685..1386

FT /*tag= e

FT /number= Ex7

CA2194256-A.

PD 05-MAR-1998.

PE 02-JAN-1997; 194256.

PR 05-SEP-1996; US-711417.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K;

DR WPI; 98-378292/33.

DR P-PSDB; W70964.

PT New nucleic acid encoding Ikaros protein involved in early

PT differentiation of lymphocytes - existing in several isoforms, and

PT related products, used to treat e.g. immune diseases or cancer and

PT to control cell differentiation

PS Claim 7; Page 70-72; 158pp; English.

CC This is the nucleotide of human Ikaros cDNA (isoform hix-1) that

CC codes for a 461-amino acid zinc finger protein (see W70964) that is

CC involved in the early differentiation of lymphocytes. It was

CC isolated from a Jurkat T cell line library using mouse Ikaros exon

CC 7 cDNA as probe. The Ikaros gene maps between pl1.2-pl3 on human

CC chromosome 7. The human and murine Ikaros sequences (see V42805-11

CC and V42840) are highly conserved. Differential splicing of Ikaros

CC gene transcripts gives rise to different Ikaros protein isoforms.
 CC The invention provides Ikaros nucleic acids, vectors and host cells
 CC expressing Ikaros proteins. These are used to treat T and B cell
 CC diseases (e.g. immune deficiencies caused by drugs, radiation or
 CC cancers), to control expression of heterologous genes placed under
 CC control of an Ikaros-responsive element, to treat nervous system
 CC diseases (e.g. Alzheimer's disease) and to modulate cell division,
 CC amplification or differentiation, especially in haematopoietic
 CC cells. Some Ikaros isoforms are antagonistic of others and may be
 CC used to inhibit interaction with DNA sequences. The same effect
 CC can be achieved with Ikaros-binding oligonucleotides. Examining
 CC the expression of the Ikaros gene, or its allelic structure, can be
 CC used to assess risk of acquiring the above diseases. 234 T;
 SQ Sequence 1386 BP; 352 A; 403 C; 397 G;

Query Match 20.8%; Score 411.8; DB 1; Length 1386;

Best Local Similarity 63.3%; Pred. No. 5.3e-94;

Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagacaacacagtggaagatgaactgcagctgtgcgggttatctctgattagatt 759
 Db 162 CATTCGACTTCCTAACGGAAACTAAAGTGTATATCTGTGGGATCATTTGCATCGGGCC 221
 Qy 760 caacgtcttgatgttcataagcgaagccataccggaagccggtccagtgtaataca 819
 Db 222 CAATGTCTCATGTTTCACAAAAGAACCCACACTGGAGAACGGCCCTTCCAGTGTCAATCA 281
 Qy 820 gtgcggggcatcttttactcagaaggttaacctctcgtcatattaaactgcacacggg 879
 Db 282 GTGCGGGGCCCTCATTCACCCAGAGGGCACTGTCTCGGCACATCAAGCTGCATTCGG 341
 Qy 880 ggaataacctttaaagtgtcaacctgtcaactacatgcacaaaggagagatgcctcac 939
 Db 342 GGAGAAGCCCTTCAAAATGCCACTCTGCAACTACGCCCTGCCCGCGGAGGACGCCCTCAC 401
 Qy 940 gggacaccttagcacacattctgtgagaagccgtacaaagtgtgaagtctcggaagaag 999
 Db 402 TGGCCACCTGAGGACGCACTCGGTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAG 461
 Qy 1000 ctacaagcagaagctccctggaggagcacaaggaacgcgtgcgcgagcttttcttcagaa 1059
 Db 462 CTATAACAGCGCAACGCTCTTTAGAGGAACATAAAGAGCGCTGCCACACTACTTGGAAAG 521
 Qy 1060 ccttgacct--gggggacgtgcgaagtgtggaggcaagacacatcaaaagc----- 1108
 Db 522 CATGGCCCTTCGGGCACTGTACCCAGTCAAAAAGAAAGAACTAAGCACAGCTGAAT 581
 Qy 1108 -----cgagatgggaagtgaagagctcgtcctctggacagattagaag 1152
 Db 582 GGCAGAACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACAGACTAGCAAG 641
 Qy 1153 caatgtgctaagcgaataagctcgtatgcctcagaataattcatcggtgagaagcgcaactg 1212
 Db 642 TAATGTGCCCAACGTAAGAGCTCTATGCCCTCAGAAATTTCTTGGGGACAAGGGCCCTGGC 701
 Qy 1213 cttcgatgccactacaatcccggtctacatgtacgagaaggaagacagatgatgcagac 1272
 Db 702 CGACACGGCCCTACGACAGTGC-----CACGTACGAGAAGGAGAGAAATGATGAAGTC 755
 Qy 1273 ccggatgaggaagcgaacatacaataagccatcagctatctatagggggtgaagccttcg 1332
 Db 756 CCACGTGTATGGACCAAGGCATCAACAACGCCATCAACTACCTGGGGCGCGAGGCCCTCGC 815
 Qy 1333 ccccttagtcagactccgctcgtcccacctctgagatggtcccagtcacagcagtg 1392
 Db 816 CCGCTGTGTGACAGCGCCCGCG---GCGGTTCAGAGTGTGTCGCGGTATCAGCCCCAT 872
 Qy 1393 gtac-----cccatagacttactctggggccgcatatgccaatggggggccccgca 1440
 Db 873 GTACCAGCTGACAGGGCGTTCGAGGGGACCCCGCGCTCCAACCACTCGGGCCAGGACAG 932
 Qy 1441 ggagatggaagaagaacggatcctcctccagagagaagattcttgcctctctgaacaggct 1500

Db 933 CGCGTGAGTACCTGCTGCTCTCAAGGCCAAGTTGTTGCTCGGAGCGGAGGC 992
Qy 1501 gtcccccaataacagtgccaggagatccacagacacagacacagacacagagatc----- 1557
Db 993 GTCCCCGAGCAACAGCTGCCAAGACTCCACGACACACCGAGAGCAACAACAGGAGCAGCG 1052
Qy 1557 --gcaacatctctaccagcaaaagcagctggtctctccccccagggcccgcaatggatgcc 1614
Db 1053 CAGCGGTCTTATCTACCTTACCTAGCAACACATCGCCGAGCGCGGACGCGTGTGCG----- 1108
Qy 1615 tctttgaagaggtccctctctcttttgaactctctcaagccccctccctctctgctgag 1674
Db 1108 ---CTCAAGGAGGAGCAGCGCGCTACGACCTGTGCGCGCGCTCCGAGAACTCGCA 1163
Qy 1675 ggaactccatcaagtgtatcaacaaagagggaggtgagtgatgtgttcttcgatgacca 1734
Db 1164 GACGCGCTCGCGTGTGTCAGCACAGCGGGGAGCAGATGAAGTGTACAAGTGCRAACA 1223
Qy 1735 ctgcccagctctctctctagattatgtatgttccaccatccacatgggtggtccatggtt 1794
Db 1224 CTGCGGGTGTCTTCTTGGATFACGCTCATGTACACCATCCACATFGGGCTGCCAGCGCTT 1283
Qy 1795 ccgtgatccctttagtgaacatgtgtggtatcgagccacacgctgctatgattctc 1854
Db 1284 CCGTGATCTTTTGAGTCAACATGTGCGGTACCCACAGCCAGGACCGGTACGAGTTCTC 1343
Qy 1855 ctctcacatccagagagagacag 1881
Db 1344 GTCCGACATAACGCGAGGGGAGCACCG 1370

RESULT 5

V66969
ID V66969 standard; cDNA; 1386 BP.
AC V66969;
DE 14-JAN-1999 (first entry)
DT Human Ikaros encoding cDNA.
KW CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW differentiation marker; immune system; corpus striatum; AIDS;
OS Alzheimer's disease; ss.
FH Homo sapiens.
FT Key
FT CDS
FT Location/Qualifiers
FT 1..1386
FT /*tag= a
FT /product= "Ikaros"
FN US5824770-A.
PD 20-OCT-1998.
PF 05-JUN-1995; 465590.
PR 02-MAY-1994; US-238212.
PR 14-SEP-1992; US-946233.
PR 14-SEP-1993; US-121438.
PR 05-JUN-1995; US-465590.
PA (GEO) GEN HOSPITAL CORP.
PI Georgopoulos K;
DR WPI: 98-582621/49.
DR P-PSDB: W72672.
PT Ikaros poly:peptide(s) - useful for treating disorders of immune
PT system or corpus striatum
PS Disclosure: Column 55-58; 11pp; English.
CC The present invention describes a purified peptide having at least one
CC of the following properties: (a) it stimulates transcription of a DNA
CC sequence under the control of a delta A element, an NFkB element or an
CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC a delta A element, an NFkB element or an Ikaros binding oligonucleotide
CC consensus sequence; (c) it competitively inhibits the binding of a
CC naturally occurring Ikaros isoform to any of a delta A element, an NFkB
CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC (e) it inhibits protein-protein interactions of transcriptional complexes
CC formed with naturally occurring Ikaros isoforms. The proteins, provided
CC that they stimulate gene transcription under the control of delta A
CC elements, NFkB elements and/or Ikaros-binding oligonucleotides, bind to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,

CC competitively inhibit binding of naturally occurring Ikaros isoforms to
CC delta A elements, NFkB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC inhibit protein-protein interactions of transcriptional complexes with
CC naturally occurring Ikaros isoforms, can be used to treat immune system
CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC Alzheimer's disease. The present sequence encodes a specifically
CC claimed human Ikaros protein.
SQ Sequence 1386 BP; 350 A; 403 C; 395 G; 238 T;

Query Match 20.8%; Score 411.8; DB 1; Length 1386;
Best Local Similarity 63.3%; Pred. No. 5.3e-94;
Matches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

Qy 700 cagcagacacacacagtggaagatgaactgcagctgtgcgggttatccctgcattagctt 759
Db 162 CATTTCGACTTCTTAACGGAACCTAAAGTGTATATCTGTGGATCATTTTCATCGGGCC 221
Qy 760 caacgtcttgatggttcataagcgaagccataccggcgaaccccgcttccagtgtaatca 819
Db 222 CAATGTGCTCTATGGTTTCAAAAAGAACCACTGGAGAAACGGCCCTTCCAGTGCATCA 281
Qy 820 gtgcggggcatcttttactcagaaagttaacctcctccgtcattataaactgcacacgg 879
Db 282 GTGCGGGGCTCATTCACCAAGAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGG 341
Qy 880 ggaacacccctttaaagtgtcacctctgcaactcgcacgctgcaaaaggagagatgcgtcac 939
Db 342 GGAGAAGCCCTTCAATGCCACCTCTGCAACTAGCCCTGCCGCGGAGGGACGCCCTCAC 401
Qy 940 gggacaccttaggacacattctgtggagagccgttacaagtgtgagttctgcggaagaag 999
Db 402 TGGCCACCTGAGGACGACCTCCCTGTTGTAACCTCAAAATGTGGATATTGTGCCCAAG 461
Qy 1000 ctacaagcagagaagctccctgcagagagcagaagcagcgtgcgagcttttcttcagaa 1059
Db 462 CTATAACAGCGAAGCTCTTTAGAGGAACATAAAGAGCGCTGCCAACACTACTTGGAAAG 521
Qy 1060 ccttgacct--gggggagcgtgcaagtgtgagggaagacacacatacgaagc----- 1108
Db 522 CATGGGCTTCCGGGACACACTGTACCCAGTCATTAAAGAGAACTAAGCACAGTGAAT 581
Qy 1108 -----cgagatgggaagtggagagctctcgtcctgcagacattagcaag 1152
Db 582 GGCAGAGAAGACTGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGCACAGACTAGCAAG 641
Qy 1153 caatgtgctaaagcaaaagctcgtgcctcagaaattcattcgtgtagaagcggaactg 1212
Db 642 TAATGTCCCAACAGTAAAGAGCTCTATGCCCTCAGAAATTTCTTGGGGACAAAGGCCCTGTC 701
Qy 1213 ctctgatgccaaactacatcccggtctacatgtacgagaaggaagacgatgatgcagac 1272
Db 702 CGACACCCCTTACGACAGTGC-----CAGTACGAGAAGGAGAACCAATGATGAAGTC 755
Qy 1273 ccggatgatggaccaagccatcaataacgccaatcagctatctagggggtgaagccttcg 1332
Db 756 CCACGTGATGACCAAGCCATCAACACGCCATCACTACCTTGGGGCCGAGTCCCTGCG 815
Qy 1333 ccccttagtcagacacgctcgtctccaccttgagatggtccagtcattcagcagtgt 1392
Db 816 CCCCTGTGTGAGACGCCCCCGG---CGGTTCGAGAGTGTGTCCTGCTATCAGCCCCGAT 872
Qy 1393 gtaccccata-----gcaacttaactcggccgcatatgcacaatgggggccccgca 1440
Db 873 GTACCACTGCACAGGCGCTCGGAGGACACCCCGCTCCCAACCACTCGGCCCCAGGACAG 932
Qy 1441 ggaagatgaaagaaacggtatcctcctccagagaagatcttgccttctgaagcaggtct 1500
Db 933 CGCGGTGGATACCTGCTGTGCTCTCTCAAGGGCAAGTGTGTGCTCGGAGCGGAGGC 992
Qy 1501 gtcccccaataacagtgccagactccacagacacagacacacacacagagatc---- 1557

CC committed lymphoid progenitors and in T and B cells, and is a
 CC transcriptional activator of a lymphoid gene. Aiolos cDNA can be
 CC used to produce recombinant Aiolos polypeptide (W15575) in host
 CC cells. The Aiolos polypeptide, coding sequence or cells expressing
 CC Aiolos may be used to treat a disorder in an animal, esp. by gene
 CC therapy. Such disorders include T-cell leukaemias and lymphomas.
 CC Non-wildtype gene structure or expression is indicative of a risk
 CC for such a disorder. Transgenic animals with an Aiolos transgene
 CC are provided.
 SQ Sequence 628 BP; 207 A; 128 C; 155 G; 138 T;

Query Match 25.0%; Score 496.8; DB 1; Length 628;
 Best Local Similarity 88.1%; Pred. No. 2.1e-115;
 Matches 553; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

Qy 569 gacagagatgagaacattatgaagccgagcccatggagatgcagagagagtgaaatg 628
 Db 1 GAAAGATGAGAAATGTTTAAAGTCAGAACCCATGGGAAATGCAAGAGCCCTGAATC 60
 Qy 629 cttacagatgcagagaatacacgagctatgaagcattaaagctggagagaca---c 685
 Db 61 CCTTACAGCTATTCAAGAGATATATGAATATGAAACATTTAAGTTGGAGAGACATGTT 120
 Qy 686 gtgcctatgacacagacagacacacacagatgggaagatgaactgcgagctgctggggtta 745
 Db 121 GTCTCATTCGATAGTAGAGAGCCACACAGTGGAAAGATGAATGGCGATGTGTGCGATTA 180
 Qy 746 tctgcattagtttcaacgtcttgatggttcataagcgaagccatccagcggaagcccg 805
 Db 181 TCTTCATCAGCTTCAATGCTTTATGTTTATAGCGNAGCCATCTGTTGAAGCCCA 240
 Qy 806 ttccagtgtaactagtgccgggcatcttttactcagaaggttaacctctcgtcatatt 865
 Db 241 TTCCAGTGTAAATCAGTGTGGGCATCTTTTACTCAGAAAGGTAACTCTCCGCCACATT 300
 Qy 866 aaactgcacagggggaaaaaccccttttaagtgtcacctctgcaactagcagcatgccaaagg 925
 Db 301 AAATGTCACACAGGGGAAACCTTTTAAAGTGTCACTCTGCAATATGTCATGCCAAAGA 360
 Qy 926 agagatgcgtcacgggacacaccttaggcacacattctgtggaagcccgacacagtgtag 985
 Db 361 AGAGATGCGCTCACGGGGCATCTTAGGACACATTTCTGTGGAAACCCCTACAAATGTGAG 420
 Qy 986 ttctgcggaagctacagcagagaagctccctggagagacacaaaggaacgtgccga 1045
 Db 421 TTTTGTGAAGAGGTATACAAGCAGAGAAGTTCCTTTGAGGACACAAGGAGCGCTGCCGT 480
 Qy 1046 gcttttttcagaacctgacctgggggagcgtgcaagtgtgagggcgaagacacataaaa 1105
 Db 481 ACATTTCTTCAGAGCACTGACCACGGGGACACTGCAAGTGGCGAGGCAAGACACATCAAA 540
 Qy 1106 gccgagatgggaagtgcagagagctctccttgagacagattagcaagcaatgtgctaa 1165
 Db 541 GCAGAGATGGGAAGTGAAGAGACTCTCGTACTTGGACAGATTAGCAGCAATGTGGCAAAA 600
 Qy 1166 cgaagaaagctcagctcctcagaaattca 1193
 Db 601 CGAAAAAGCTCAATGCTCAGAAATTCA 628

RESULT 3
 ID Q44980
 AC Q44980;
 DT 21-OCT-1994 (first entry)
 DE Human Ikaros peptide coding sequence.
 KW Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
 OS corpus striatum; regulatory gene; ss.
 FH Homo sapiens.
 FT Key
 FT cds Location/Qualifiers
 1..1611
 /*tag= a

FT W09406814-A.
 PD 31-MAR-1994.
 PF 14-SEP-1993; U08743.
 PR 14-SEP-1992; US-946233.
 PA (GEO) GEN HOSPITAL CORP.
 PI Georgopoulos K;
 DR WPI; 94-118387/14.
 DR P-PSDB; R46964.
 FT T-cell pathway regulatory gene, Ikaros - encodes family of unique
 FT zinc finger proteins, useful for treating immune system disorders
 PS Claim 13; Page 44-46; 112pp; English.
 CC The Ikaros gene encodes a zinc finger protein which can be used in a
 CC therapeutic composition to treat animals with an immune system
 CC disorder. It may also be used for assessing whether a subject is at
 CC risk for an immune disorder. It is of particular use in treating a
 CC disorder of the corpus striatum.
 CC Sequence 1611 BP; 375 A; 484 C; 480 G; 272 T;

/product= Peptide with Ikaros activity.
 Query Match 20.9%; Score 413.8; DB 1; Length 1611;
 Best Local Similarity 63.1%; Pred. No. 1.8e-94;
 Matches 785; Conservative 0; Mismatches 397; Indels 63; Gaps 7;
 Qy 700 cagcagacacacacagtggaagatgaactgcgacgtgtgcggttatcctgcattagctt 759
 Db 384 CATTCGACTTCTTACCGGAAACATAAAGTGTGATATCTGTGGGATCATTTTCATCGGGCC 443
 Qy 760 caacgtctgtatgttcataagcgaacccatccgcgaacgcccgttccagtgtaatca 819
 Db 444 CAATGTGCTCATGTTTACAAAAGACCCACTGGAGAACGGCCCTTCCAGTGCAATCA 503
 Qy 820 gtgcgggacatcttttactcagaaggttaacctctcgtcatattaaactgcacaggg 879
 Db 504 GTGGGGGSCCTCATTTACCCAGAAAGGGCAACCTGCTCCGGCACATCAAGCTGCTTCGG 563
 Qy 880 ggaacaccccttttaagtgtcacctctgcaactcgcagatgccaaaggagagatgcgtcac 939
 Db 564 GGAGAAGCCCTTCAATAGCCACCTCTGCACTAGCCCTGCGCGGAGGGACGCCCTCAC 623
 Qy 940 gggacaccttaggcacattctgtggagaagccgtacaaagtgtgagttcttcggaagaag 999
 Db 624 TGGCCACTGTAGGAGCGACACTCCGTTGGTAAACCTCACAATGTGTGATATTGTGCCGAAG 683
 Qy 1000 ctacaagcagagaagctccctggagagacaaaggaacgctgcgagagcttttcttcagaa 1059
 Db 684 CTATAACAGCGAAGCTTTTAGAGGAACATAAAGAGCGCTGCCACAACACTACTTGGAAAG 743
 Qy 1060 ccttgacct--ggggagcgtgcgaagtgtggaggaagacacacataaagc----- 1108
 Db 744 CATGGGCTTCCGGGCACACTGTACCCAGTCAATTAAGAAGAAACATAAGCACAGTGAAT 803
 Qy 1108 -----cagagatggaagtgcagagctctcgtctcgtgcgacagattagcaag 1152
 Db 804 GGCAGAGACACTGTGGAAGATAGGATCAGAGAGATCTCTCGTCTGGACACACTAGCAAG 863
 Qy 1153 caatgtggttaagcgaagaaagctgcgtccagaaattcatcgtgtgagaagcggcactg 1212
 Db 864 TAATGTGCGCAACGTAAGAGCTCTATGCTTCAGAAATTTCTTGGGACACAGGGCCCTGTC 923
 Qy 1213 ctctgatgcgaactacatacccggtctacatgtacagaagagagacagatgtagcagac 1272
 Db 924 CGACACGCCCTTACGACAGTGC-----CACGTAGGAGAAGGAGAACAATGATGAAGTC 977
 Qy 1273 ccggaatgagccaagccatcaataagccatcagctatctataggggtgagccttcgc 1332
 Db 978 CCAGGTGATGACCAAGCCATCAACAACGCCCTCACTACCTTGGGGCCGAGTCCCTGCG 1037
 Qy 1333 ccccttagtcagactccgctctcccaacctgtgagatgtccagtcacacagcagtg 1392
 Db 1038 CCCGCTGTGTCAGACGCCCGCGG---CGGTTCGAGGTGTCCCGGTCTCATCAGCCGAT 1094
 Qy 1393 gtaccccaata-----gcacttaactcggggcgagatatgccaatggggggcccgca 1440

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 1999, 21:32:29 ; Search time 70.34 Seconds
(without alignments)
7056.882 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagcagcgacacgcgtctcg.....gaactcaaacaccctcgcag 1984

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980.8	99.8	1984	1	Mouse Aiolos cDNA.
2	436.8	25.0	628	1	Human Aiolos parti
3	413.8	20.9	1611	1	Human Ikaros pepti
4	411.8	20.8	1386	1	Human Ikaros cDNA
5	411.8	20.8	1386	1	Human Ikaros encod
6	411.8	20.8	1386	1	Human Ikaros isofo
7	411.8	20.8	1551	1	Human Ikaros isofo
8	379.6	19.1	2049	1	Murine Ikaros cDNA
9	379.6	19.1	2049	1	Mouse Ikaros encod
10	379.6	19.1	2049	1	Mouse Ikaros isofo
11	348	17.5	1788	1	Mouse Ikaros isofo
12	346.4	17.5	1788	1	Murine Ikaros gene
13	346.4	17.5	1788	1	Murine Ikaros cDNA
14	346.4	17.5	1788	1	Murine Ikaros enc
15	269.6	13.6	1004	1	Ikaros cDNA. Trans
16	269.6	13.6	1004	1	Ikaros isoform enc
17	269.6	13.6	1004	1	Human Ikaros cDNA.
18	236.8	11.9	1170	1	Mouse Ikaros isofo
19	235.2	11.9	1170	1	Murine Ikaros cDNA
20	235.2	11.9	1170	1	Mouse Ikaros encod
21	200.4	10.1	714	1	Ikaros protein enc
22	191.2	9.6	1296	1	Murine Ikaros cDNA
23	191.2	9.6	1128	1	Murine Ikaros cDNA
24	191.2	9.6	708	1	Ikaros protein enc
25	191.2	9.6	1128	1	Mouse Ikaros encod
26	191.2	9.6	1296	1	Mouse Ikaros encod
27	191.2	9.6	1296	1	Mouse Ikaros isofo
28	191.2	9.6	1128	1	Mouse Ikaros isofo
29	109.4	5.5	168	1	Ikaros protein enc
30	104.6	5.3	168	1	Ikaros protein enc
31	73	3.7	1663	1	Renal cancer assoc
32	70.8	3.6	1683	1	Human SRE-2BP anal
33	70.8	3.6	2168	1	Human SRE-2BP anal
34	70	3.5	353	1	EST clone AW95. Ne
35	66.4	3.3	264	1	DNA encoding a nuc
36	66.4	3.3	264	1	DNA encoding an an
37	66.4	3.3	264	1	Zinc finger protei
38	66.2	3.3	2680	1	Myc-binding zinc-f
39	64.8	3.3	1476	1	Zinc finger protei
40	64.4	3.2	323	1	Human gene signatu
41	63	3.2	2133	1	Human OTK18 gene.
42	63	3.2	3754	1	Human OTK18 gene.
43	62.8	3.2	270	1	EST clone CO1069.

44 61.4 3.1 3776 1 T42903 TRP-1 protein codi
45 60.2 3.0 3720 1 Q80513 Genetic locus bcl-

ALIGNMENTS

RESULT 1

T60490
ID T60490 standard; cDNA; 1984 BP.

AC T60490;

DT 07-JUL-1997 (first entry)

DE Mouse Aiolos cDNA.

KW Aiolos; transcription activator; immune system; T lymphocyte;

KW B lymphocyte; leukaemia; lymphoma; asthma; gene therapy;

KW transgenic animal; ss.

OS Mus sp.

FH key Location/Qualifiers

FT cds 374..1897

FN M0371471A-AJ. /*tag= a

PD 24-APR-1997.

PF 17-OCT-1996; U16774.

PR 18-OCT-1995; US-005529.

PR 14-MAY-1996; US-017646.

PA (GEO) GEN HOSPITAL CORP.

PI Georgopoulos K, Morgan BA;

DR WPI; 97-245047/22.

DR P-PSDB; W15574.

PT Aiolos polypeptide and corresponding DNA - used to reconstitute a

PT mammalian immune system, for the treatment of T cell leukaemia(s)

PS and lymphoma(s)

CC A cDNA clone (T60490) corresponds to the mouse Aiolos gene, a

CC homologue of Ikaros whose expression is restricted to lymphoid

CC lineage. Aiolos cDNA was isolated from a mouse spleen cDNA library

CC using a probe that spanned nucleotides 796-1156 of the sequence.

CC Aiolos (W15574) can form dimers with Aiolos or Ikaros polypeptides,

CC is expressed in committed lymphoid progenitors and in T and B cells,

CC and is a transcriptional activator of a lymphoid gene. Primers

CC (see also T60494-95) based on mouse Aiolos gene exons were used to

CC amplify a human partial Aiolos cDNA (see also T60491). Aiolos cDNA

CC can be used to produce recombinant Aiolos in transformed host cells.

CC The Aiolos polypeptide, coding sequence or cells expressing Aiolos

CC may be used to treat a disorder in an animal, esp. by gene therapy.

CC Such disorders include T-cell leukaemias and lymphomas. Transgenic

CC animals are provided with an Aiolos transgene.

CC Sequence 1984 BP; 512 A; 570 C; 535 G; 367 T;

Sequence 1984 BP; 512 A; 570 C; 535 G; 367 T;

Query Match 99.8%; Score 1980.8; DB 1; Length 1984;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 cagcagcgacacgcgtctcgtctcgttcgacacgcctcattcccggtgtttctcaag 60

Db 1 CACGAGCGCACACGGTCGGCTCTCTTGGACACGCCCTCATCCCGGTGTTTCTCAAG 60

Qy 61 tagacgtcccgagacggtcgtcgttcgacgtcttccacgcatcagggttcctcaggctt 120

Db 61 TAGACGTCCTCCGAGACGGTCGCTGAGGCACCTGTTTCCACGGATCAGGGTTCTCAGGCTT 120

Qy 121 gacattcaaaagtgggtgcggaaaccccgcgacactcggagcgtgtgtttaaagcgccgcca 180

Db 121 GACATTCAAAAGTGGGTGCGGACACCCGCGCACTCGGAGCGTGTCTTTAAAGCGGCCCA 180

Qy 181 gccagcgccgctctaaactcgcgcccggtgcgcggcggtcccgccctcatctgccc 240

Db 181 GCCAGCGCCGCTCTAACTCGCGCCCGGCTGCGGGGCTCCCGCCTCATCTGCGCC 240

Qy 241 gacgacccgagcgtatcccgggacccctccctgcgccggaatctcccgccgagcgcg 300

Db 241 GACGCGACCGAGCGATCCCCGGGGCTCTCCCTGCGCCCGGAATCTCCCGCCAGCGCGCGG 300

YKORSLIEHKKHCHNYLONVMEAGGIVHHASPMDECKNPETMMNNLSMPERP
AVTERLASNKKRSKSTPPQRFELGKLMRYGYPDLHFDMAEYKEAEILIQSOMDQAINN
AIIYLGDADLRPLIHHSAAAMPEVPIIVSSLYSQVYHPARVPTRETSDSNDNNMD
GPISLIRPNHPQERASPSNCLDITDESSEHAPAILPGNHALNSKRKQSSVYPKE
HGKPLDATKSSLSGSKDVYRVFNSEGEQIRAFKCEHRCRVFLDHNVTTHMG"

BASE COUNT 295 a 233 c 244 g 214 t
ORIGIN

Query Match 7.4%; Score 146.6; DB 4; Length 986;
Best Local Similarity 52.0%; Pred. No. 2.1e-23;
Matches 508; Conservative 0; Mismatches 394; Indels 75; Gaps 5;

Qy	878	ggggaaaaaccccttttaagtgtcacctctgcaactacatgcagccaaaggagagatgcctc	937
Db	12	GGGGAGAGCGGTTAAATCTTCATCTTCAGCTATGCATGTCGGAGAGAGACGCACTT	71
Qy	938	acgggacaccttagacacattctgtgagagccgtacaaagtgtgattctgcggaaga	997
Db	72	ACTGCCACCTCAACACACTCCGTTGGCAACCTCACAAGTGCAACTACTGTGGTCGA	131
Qy	998	aactacaagcagaagaagctccctggagagcacaaagcgtcccgagcttttcttcag	1057
Db	132	AGCTACAAGCAGAGAGAGCTCTTTGGAGCACACAAGGAGAGGTGTCACAACTACTTCAA	191
Qy	1058	aacctgacctgggggagcgtgaagtgtgg-----	1089
Db	192	AATTTGGGTGGAGGCTGGCCAAAGGATTGTGCACCATGCATCTCTATGGACGAGTGT	251
Qy	1089	-----aggaagcacatcaaacgagagatgggaagtggagagagctctc	1132
Db	252	AAGAAATCCAGAGACCATGATGGAACAAACACATGCTCTAATGCCATTTGAGAGACCGCT	311
Qy	1133	gtcctgacagattagcaagcaattgtggcgaagcgaagcgtcgatgcctcagaaattc	1192
Db	312	GTTATAGAGACTTGCAGCAACATGGGAAGCGTAAGAGCTCCACCCACAGAGGTTT	371
Qy	1193	atcgtgtgaagaagcgcactgcttgcg---atccaaactacaatcccggtacatgacgag	1249
Db	372	TTGGGTGAGAAGCTCATGAGATATGGATATCTGACTTGCACTTTGATATGGCCATATGAG	431
Qy	1250	aaggagacgagatgacgaccccgatgatggaccaagcacaataaagcccatcagc	1309
Db	432	AAAGGCGAGAAATATCCAGTCAAAATGATGATCAGGGGATTACAAATGCTATAACA	491
Qy	1310	tatctaggggtgaagccttcgcccccttagtcagactccgctgctccacactctgag	1369
Db	492	TACCTGGAGCAGATGCACCTCGGCTCTCATACCACTCAGCAGCTGCCATGCCCTGAG	551
Qy	1370	atggtcccaagtacagcagtggttaccccaatagcaacttaactcgggc-----	1418
Db	552	GTTCCACCAATTTGTCAGTTCACTTTACTCTCAGGTCTATCATCCAGCCAGAGTGGAGAGA	611
Qy	1418	-----gatagccaaatggggcccgagagatggaaagaaagcagatcctc	1465
Db	612	CCCACAGCAGGAGACATCAGACAGACACGATACACACATGGATGGCCCAATCTCCCTT	671
Qy	1466	ctgcagagaagattgtcccttctgaacgaggtgtgtcccccataaacagtgccccagac	1525
Db	672	ATCAGACCAAGAACACCCTCAGGAGAGGGAGGCATCACCCAGCAACAGCTGCTTGAT	731
Qy	1526	tccacagacaccgacagcaaccagagagatggccaacatctctaccaggaagccacgtg	1585
Db	732	ACTACTGACTCAGAAAGACGACCATGAGG--CACCGGCAATCTTACCAGGCAATCATGCCT	789
Qy	1586	gtcctccccagcccgcaatgggagtcctctctgaagagaggtccctcgctcttttga	1645
Db	790	TAAATTCAAAGAGAAAGCAAGATTGGGTGATCCCAAGGAACATGGCAAC-CATTGGAT	848
Qy	1646	ctcctcaagcccccctcccatctgctgagggactccatccaaagtgatcaacaagaaggg	1705
Db	849	GCCACAAAGTCCAGCTTGGGTTCTCCAAAGGATGCTACAGAGTTTTTAATAGTGAAGGA	908

Qy	1706	gaggtgatggatgtgttttgatgtgaccactgcccagtcctcttctctagattatgtgatg	1765
Db	909	GAGCAGATTAGGCGCATTTAAATGTGAGCACTGTCCGGTTCTGTTCTTGACCATGTTATG	968
Qy	1766	ttcaccatccacatggg	1782
Db	969	TATACGATCCATATGGG	985

Search completed: November 6, 1999, 06:39:34
Job time: 4688 sec

Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
Location/Qualifiers
1. .2079
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
81..1232
/genes="Ikaros"
81..1232
/genes="Ikaros"
/notes="Ik-8 isoform; similar to mouse and human Ikaros/LyF-1; alternatively spliced form missing exons 3 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53433.1"
/db_xref="PID:g2062738"
/db_xref="GI:2062738"
/translation="MEMERAQMSQMPGRDPPNDVSENDAMPIDPLSASNLQ HNRGKEGLGERPFQTCGASFTQKGNLLRHILKHSGEKFKCHLCNYACRRDAL SGLRTHSVGKPKHCAKCYGRYSKORSSLEHKEKCHNYLQCMGLQNSIYTGKRFNSI SPEGPELMQPHVIDOAINSAINYLGAESLRLIQTSPSSDMVMGSMYPLHRPPA EGHGLSAKSAANELLALKSASKEKDGSPSHGSDSTDTESNEEKAGVAGASGLI YLTHHTSCVRNGLPLVKEOQRYEAMRASIEIETASEGFKVLSGEQVRYRCEHC RILFLDHVMTIHMCGHGRDFEFCNCHRSQDRFESSHMRGHEHRI"
BASE COUNT 593 a 500 c 464 g 521 t 1 others
ORIGIN

Query Match 8.3%; Score 164; DB 4; Length 2079;
Best Local Similarity 52.4%; Pred. No. 2.6e-27;
Matches 579; Conservative 0; Mismatches 380; Indels 147; Gaps 4;

QY 794 ggcgaacgcccgttcagtgtaacatcagtcgaggcgatctttttactcagaaggttaacctc 853
DB 243 GGAGAGGTCATTCAGTGCACCCAGTGTGGCGTCTTTTACCAGAGGCAACCTG 302

QY 854 ctccgtcaataaactgcacacgagggggaaacaccttttaagtgtcacctctgcaactac 913
DB 303 CTACGTACATCAAGTCCACTCAGGAGAGAGCCCTTCAAGTGTCACTGTGCAACTAT 362

QY 914 gcataagcgaagagatgcctcagcgacaccttagacacatctgtggagagccg 973
DB 363 GCTGTGCGGAGAGAGCGCCCTCAGCGGTCACTGCGTACCCACTCTGTGGAAACCC 422

QY 974 tacaagtgtgagttctcggaaagatacaagcagagaagctccctggaggagcacaag 1033
DB 423 CACAAGTGTCTTACTGTGGCGGAGCTACAGCAGCGTAGCTCTGTGGAGAACACAAG 482

QY 1034 gaacgtgcgagctttcttcagaacacctgacctgggggacgctgcaagtgtggagca 1093
DB 483 GAGCGGTGTACAACTACCTCCAG----- 507

QY 1094 agacacataagcagagatgggaagtggagagctctcgtcctggacagattagcaagc 1153
DB 507 ----- 507

QY 1154 aatgtggtaagcgaaaaagctgatgcctcagaataattcatcgtgtgagacggaactgc 1213
DB 507 -----TGCATGGGCTCCAGAACAGCATCTATACAGGTGAGAAACGC 548

QY 1214 ttgatgcacactacaatcccgctacatgtacagagaagagacagatgtatgcagacc 1273
DB 549 TTCTCCAACTATCTCTTCGAGGTGGCCCGGGGGAAC-----TCATGCAGCCT 596

QY 1274 cggatgtgacacaaagcattacaacgcatcagctatctataggcgctgaagccttcgcg 1333
DB 597 CATGTCTATCACCAGGCCATCAACAGTGCCTCACTACTCTGGTGGCCGAGTCACTCCGA 656

QY 1334 cccttagtcagactccgctgtctccacacctgtgagatggtccacagtcacacagcagtg 1393
DB 657 CCGCTCATCCAGACCTCCCCAACCTCCTCTGACATGGGGGTGATGGGCTCCATGTACCCC 716

QY 1394 taccocatagcacttactctggtggccgatatgccaatggggggcccgagagagatgaaaaag 1453
DB 717 CTCCAAGCCCTTCAGAGGGCCACGGCTGTGAGCAAGGACAGCAGCCGAAAAT 776

QY 1454 aaacgatctctccagagagaatcttgccttctgaacgaggtctgtcccccataac 1513
DB 777 CTTCTGTCTCGAAAGTCCAAGTCAGCTCCAGGAGAGACGGCTTCCCCAGCCAC 836

QY 1514 agtgcacagactccacagacacccagacacacacacagaga-----tcgc 1558
DB 837 AGTGGCCAGATTCCACCCACACGGAGAGCAACAATGAGAGAAGCAGGTGTCGGGGCC 896

QY 1559 caacatctctacagcaaacacagctgttctcccccagggccgcaatgggatgctctt 1618
DB 897 TCAGGGCTCATCTACCTGACAAACACATCAGCTCCGGGTACGTAATGGCTGTGCTGCC 956

QY 1619 ct-----gaagaggttcctcgtctttgaactcctcaagccccctccatctgcctg 1672
DB 957 CTGGTGAAGGAGGAGCAGCAGCGGAGTACGAGGCCATCGGGGCCAGTATCAGATTGCC 1016

QY 1673 agggactccatcaaatgatcatcaaaaagaggggaggtgatgtgttttcgatgtgac 1732
DB 1017 TCGGAGGGTTTCAAGGTGCTGAGTGAGAGGGGAGCAGGTGAGGCATATCGCTGTGAA 1076

QY 1733 cactgcacgtct 1792
DB 1077 CACTGCGGCATCTCTCTCTAGACCATGTGTATACACCATCCAGTGGGTGCGCATGCG 1136

QY 1793 ttccgtgatcccttgatgtaacatgtgtgctcatcaagcagcagatcgtatagttc 1852
DB 1137 TTCGAGAGCCCTTCGAGTGCAACCTCTGTGCCATCGCAGTCAGACCGTACGAGTTT 1196

QY 1853 tctctccatcgcgcagagagagca 1878,
DB 1197 TCGTCACATGACCCGAGGGGAGCA 1222

RESULT 15
AF024439
LOCUS AF024439 - 986 bp mRNA VRT 20-OCT-1997
DEFINITION Xenopus laevis ikaros-related transcription factor mRNA, partial cds
ACCESSION AF024439
NID 92547231
VERSION AF024439.1 GI:2547231
KEYWORDS African clawed frog.
SOURCE Xenopus laevis
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 986)
AUTHORS Turpen,J., Kelley,C., Mead,P. and Zon,L.
TITLE Bi-Potential Primitive-Definitive Hematopoietic Progenitors in the Vertebrate Embryo
JOURNAL Immunity (1997) In press
REFERENCE 2 (bases 1 to 986)
AUTHORS Turpen,J., Kelley,C., Mead,P. and Zon,L.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1997) Hematology, Children's Hospital, 300 Longwood Ave, Boston, MA 02115, USA
FEATURES
source location/Qualifiers
1. .986
/organism="Xenopus laevis"
/db_xref="taxon:8355"
<1..>986
/codon_start=3
/product="Ikaros-related transcription factor"
/protein_id="AAB81280.1"
/db_xref="PID:g2547232"
/db_xref="GI:2547232"
/translation="IRLGEKPFKPCFCSYACRRRDALTGHLRTHSVGKPKHCNYCGRS

```

Query Match      9.4%; Score 187.2; DB 12; Length 684;
Best Local Similarity 65.1%; Pred. No. 1.5e-32;
Matches 336; Conservative 0; Mismatches 153; Indels 27; Gaps 3;

QY 700 cagcagaccaccagtggaagatgaactgcagcgtgctgaggttatctcattagctt 759
DB 168 CATTCGACTCTTCTAACGGAAACTAAAGTGTATCTGTGGGATCGTTGCATCGGGCC 227
QY 760 caacgtcttgatggttcataagcgaagcattacccggaagcccgctccagtgtaatca 819
DB 228 CAATGTGCTCATGTTTCACAAAGAAAGTCATACT---GAAAGGCGCTTCCAGTGCACCA 284
QY 820 gtgcggggcattctttactcagaaggttaacctctcctcgtcatattaaactgcacacggg 879
DB 285 GTGTGGGGCCCTCTTACCCAGAAAGGCAAGCTCTCTGGCGCACATCAAGCTGCACCTCGG 344
QY 880 ggaataacaccttttaagtctacctctgcaactagcagcagcgaagagagatgcgtcac 939
DB 345 TCAGAAGCCCTTCAATGCCATCTTTGCAACTATGCTGCCGCCGGAGGAGCGCCCTCAC 404
QY 940 gggacaccttaggacacattctgtgagaagccgtacaaagtgtgagttcttcggaagaag 999
DB 405 CGGCCACCTTAGGAGCGACTCCGTTGGTAAAGCCTCACAAATGTGGATATTGTGCCGGAG 464
QY 1000 ctacaagcagagaagctccctggagagcacaaggaagcgtgcgagcttttcttcagaa 1059
DB 465 CTATAACAGGGAAGCTCTTTAGAGGAGCATAAAGAGCGGATGCCACAACACTACTTGGAAAG 524
QY 1060 cccgtgacctgggg-----acgtgcgaagtgtggaggcaagacacataaagc----- 1108
DB 525 CATGGCCCTCCGGGCATGTACCCAGTCTATTAAAGGAAGAACTAACCCACAGAGATGGC 584
QY 1108 -----ccagatgggaagtgcagagcgtctcctgcagacagattagcaagcaa 1155
DB 585 AGAAGACCTGTGCATAGTAGAGCAGAGAGGTCCCTTCTCTGGACAGCTGGCGCAAGCAA 644
QY 1156 tgtggttaagcgaataaagctcgtgcctcagaaatt 1191
DB 645 TGTCCCAACGTAAGAGCTCTATGCCTCAGAAATT 680

RESULT 13
XLU92202      415 bp      mRNA      VRT      02-MAY-1997
LOCUS      Xenopus laevis/gillii Ikaros homolog (Ikaros) mRNA, partial cds.
DEFINITION
ACCESSION      U92202
NID      g2062743
VERSION      U92202.1 GI:2062743
KEYWORDS
SOURCE      Xenopus laevis/gillii.
ORGANISM      Xenopus laevis/gillii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
Xenopus.
REFERENCE      1 (bases 1 to 415)
AUTHORS      Hansen, J.D.
TITLE      Isolation and characterization of Ikaros homologues in the rainbow
JOURNAL      trout
REFERENCE      2 (bases 1 to 415)
AUTHORS      Hansen, J.D.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
FEATURES      Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
SOURCE      Location/Qualifiers
            1..415
            /organism="Xenopus laevis/gillii"
            /db_xref="taxon:8359"
            /tissue_type="thymus"
            /clone_lib="LG15.0"
            <1..>415

gene

```

```

/gene="Ikaros"
<1..>415
/gene="Ikaros"
/note="similar to mouse and human Ikaros/LyF-1; composed
of exons 4, 5 and 6"
/codon_start=1
/product="Ikaros homolog"
/protein_id="AAB53435.1"
/db_xref="pid:g2062744"
/db_xref="gi:2062744"
/translation="GERPFCNOCGASFTQKGNLLRHILKSGEKPKCHMCNACRR
RDALSGHLRTHSVGKPKGCGRSYKQRSSLEEHKRCRCHNYLQSLGLOGLHVAKEE
SKIDMAEDLSKIGSERSLVLDRLASNAKRKSSMPQ"
BASE COUNT      124 a      93 c      100 g      98 t
ORIGIN

Query Match      8.4%; Score 166.2; DB 4; Length 415;
Best Local Similarity 65.3%; Pred. No. 7.7e-28;
Matches 271; Conservative 0; Mismatches 123; Indels 21; Gaps 1;

QY 794 ggcgaacgcccgttccagtgtaatacagtcggtgggcatcttttactcagaaggttaacctc 853
DB 1 GGTGAACGGCCTTTCAGTGCACACCACTGTGGTGCCTCTTACCCAGAGGGCAATCTC 60
QY 854 ctccgtcatattaaactgcacacgggggaaacaccttttaagtgtcacctctgcacattac 913
DB 61 CTTGCTCATATCAAAATACACTCTGTGTAAGAGCCCTTTAAATGTACATGTGTAACTAT 120
QY 914 gcatgccaaaggagatgcgtcacgcgggacacctttaggacacattctgtggagaagccg 973
DB 121 GCATGACAGGCGAGGATGACCTCAGTGGTCACTCAGGACACATCTGTAGGCAAAACCT 180
QY 974 tacaagtgtgagttcttcggaagaagctacaagcagaagcctccctggagagcacaag 1033
DB 181 CACAAGTGTGATACTGTGTGCTAGTATAGCAACGAGGAGCTCTTAGAGAGCATATAA 240
QY 1034 gaacgtgcgcgagcttttcttcagaaacctgaacctgggggacgctgcaagtgtggaggca 1093
DB 241 GAGAGATGTCACAACTACCTACAAAGTATAGGCTTGCGGGCCATCTTTATGCTGTAAAA 300
QY 1094 agacacatcaaa-----gccagatgggaagtgcagagctctc 1132
DB 301 GAAGAGAGCAAAATGATATGCGCAGAAAGACCTGTCCAAGATTTGGGTGAGAGATCACTT 360
QY 1133 gtcttgacagattagcaagcaatgtgcttaagcgaaaaagctcgtgcctcaga 1187
DB 361 GTGCTTCACAGACTTGCAGTAAGCTTGCCTTGAAGAGCTCTATGCCTCAGA 415

RESULT 14
OMU92199      2079 bp      mRNA      VRT      03-MAY-1997
LOCUS      Oncochrynychus mykiss Ikaros homolog (Ikaros) mRNA, IX-8 isoform,
DEFINITION      complete cds.
ACCESSION      U92199
NID      g2062737
VERSION      U92199.1 GI:2062737
KEYWORDS      rainbow trout.
SOURCE      Oncochrynychus mykiss
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE      1 (bases 1 to 2079)
AUTHORS      Hansen, J.D.
TITLE      Isolation and characterization of Ikaros homologues in the rainbow
JOURNAL      trout
REFERENCE      2 (bases 1 to 2079)
AUTHORS      Hansen, J.D.
TITLE      Direct Submission
JOURNAL      Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for

```

SFTQKGNLLRHIKLSHGEKPFKCHLCNYACRRRDALSGHLRTHSVGKPHKCAYCGRSY

Query Match 13.1%; Score 260.6; DB 4; Length 2309;
 Best Local Similarity 56.0%; Pred. No. 4.7e-49;
 Matches 718; Conservative 0; Mismatches 444; Indels 121; Gaps 7;

QY	700	cagcagacaaacccagtggaagatgaactggaagcgtgctgggttatctctcattagctt	759
DB	430	CATCCGCCCTCCCAACGGGAAGCTGAAGTGGATATCTGTGGGATGATTTCATTTGGGCC	489
QY	760	caacgtcttgatggtttcataaagcgaagccataacc-----	794
DB	490	CAATGTGTTGATGTTTCACAACGGAGTCACACTGAAGAAAGAAAGTCAGTTTTGGRAACA	549
QY	794	-----ggcgaacgcccggttccagtgtaactgacgtgcgggcatcttttaactcagaag	846
DB	550	ACAAAAGGTGAAGAGCCATTCCAGTGAACATCAATGTGCTTCATTCACCTCAGAAAGG	609
QY	847	taacctctccctcatattaaactcacacggtgggaaacacctttaaagttaacctctg	906
DB	610	TAACCTGCTCCGACACATCAAACTTCACTCTGGCGAGAAACCTTTCAAAATGTCACCTGTG	669
QY	907	caactacgcatgccaaaggagatgcgtccacgggacaccttaggacacattctgtgga	966
DB	670	CACATATGCTTCCGCCCGCAGAGAGCTCTCACTGGACATCTGCCACATCATTCGGTTGG	729
QY	967	gaagccgtacaagtgtgattctgcgggaagctacagcagagaagctccctggaggga	1026
DB	730	AAAGCCCCATAAGTGTGATATTGGGAGCGCAGTTTACAAGCAGCGAGCTCACTGGAGGA	789
QY	1027	gcacaaggaaactgccaggtttcttcagaacacctgacctgggggacgtgcaagtgt	1086
DB	790	ACATAAGGAGAGATGTCAACACTTTCAGTGTGATGGCCCTTCAGAAACA--GCATTAT	847
QY	1087	ggaggcaagcacatcaaaagccgagatgggaag-----tgagaga	1126
DB	848	ACAGTGAAGGAGAGAACAGCCAGATGAGCAGAGGAGGACATGCCTGCATCTGAGAGG	907
QY	1127	gtctctgtctggacagattagcaagcaatgtggctagcgaagaaagctcgtacctcag	1186
DB	908	GCCTTGGTCTAGACAGGATAGCTAACAATGTAGTACAGCTGAAGAGCTCTATGCCACAG	967
QY	1187	aaattcatcgttgagaagcggcactgttcgatgccaaactacaatcccggtacatgtac	1246
DB	968	AGTTTGTGGGAGAGAAATGCTGTGCAGAGCTATCTTTCGAGAGTGGCTCAGGT-----	1022
QY	1247	gagaaggagaacgagatgatgcagacgagatgtggtgaccgaagccatcaataaagccatc	1306
DB	1022	-----GACTGTATGACGAGCCCATGTGATGTATCAGGCCATCAACAGTGCATTT	1069
QY	1307	agctatctaggggctgaagccttcggccccccttagtccagagactccgctgctccacacct	1366
DB	1070	AGTATATCTGGGTGCAGAGTCTTGGCGCTCTGTGTTTCAGACCTCTCTCTGGTCCGCCGAC	1129
QY	1367	gagatggtcccaagtcatcagcagatgtgtaccccatagcacttactcggcgcatatgccca	1426
DB	1130	ATGGTGGTTCAGGCCCTCTATACAACTGTCAACAGTGCACAAACAGCTGAAGGCAATGGCGTT	1189
QY	1427	atggggggcccgagagatggaagaagacggatccctcctgccagagaagatcttgcct	1486
DB	1190	TCGTCTAAGACAGCGCCGACAGGACCTTCTTCTTACTCTCTAAGTCCAAATCCGCCTCT	1249
QY	1487	tctgaacgaggtctgtcccccaataacagtgccccaggtactccacagacccagcagaac	1546
DB	1250	GTTGACAAAGACGGTTCCCCCGAGTCCAGCGGGCAGGATTCCACTGCACACTGAGAGCAAC	1309
QY	1547	cacgaagatgcgcaa-----catctctaccag	1573
DB	1310	AATGAGGAGCGGTTACGCCGGGTGAAGCGGAACAGCAGCCACAGGTTGGTCTCATCTACCTG	1369
QY	1574	caaaagccagtggtctctccccccagcccgcaatggatgcctctctcgaagaggtccct	1633
DB	1370	ACCAACCATATGGCTCCAGGTATGAAATGGAGGCCTGCCAGGGGTGAAGAGAAACAA	1429
QY	1634	---cgctcttttgaacctctcaagcccccctccatctgcctgag-----ggac	1678

```

/product="ikaros homolog"
/protein_id="AAB53434.1"
/db_xref="PID:g2062740"
/translaton="MEMEQAQMSOMPRDPPNDVSEENDEAMPIDELSSANLQ
HNNGDREGELERFOCTOGKGNLLRHKLHSGKPKFKHLNACRRRAL
SGHLRTHSVKPKHCAKCGRSYKORSLEEKHERCHNYLOCMGLONSIIYVKEESNQ
NEQEDLSAQMGRKALVLDRLANNVARKSTMPQKFGVERFNSIFEGGPELMOPH
VIDQAINSNLYGAE32PLIOTSPSSDMVNGSMYPLKPPAEHGSLKSDSAE
NLLAKRSKASSEKDSHSGQSDSTESNEEKAGVAGSLIYTNHITGVRNG
VLPLVKEQOQRQVEMARASTIEIASGEKFLVSGEEOVRAVRCHECHRLFLDHVMYTH
MGCHGRDPFECNLCGRHSODRYEFSSHMTRGHRH"
BASE COUNT      638 a   517 c   488 g   539 t      1 others
ORIGIN

Query Match      14.6%; Score 290.6; DB 4; Length 2183;
Best Local Similarity 57.2%; Pred. No. 8.2e-56;
Matches 648; Conservative 0; Mismatches 419; Indels 66; Gaps 4;

Qy 794 ggcgaaccccgcttcagatgaatcagtcgaggggacatctttactcagaagaagtaacctc 853
Db 212 GGAGAGCGTCCAGTCCAGTGCACCCAGGTGGCGCTTCTTCCACCACCAAGGCAACCTG 271

Qy 854 ctccgctaatataactcacacgggggaaacaccccttttaagtgtcacctctgcaactac 913
Db 272 CTAGGTACATCAAGTCCATCGAGGAGAGAGAGCCCTTCAAGTGTCACTGTGCAACTAT 331

Qy 914 gcatcgaagagagatgcgtcacgggacaccttagcacacattctgtgagaaagcg 973
Db 332 GCCTGTCCCGGAGAGAGCCCTCAGCGGTACCTGCGTACCACCTCTGTGGAAACCC 391

Qy 974 tacaagtgagttctcgggaagagctacaaagcagaagaagctccctggaggagcacaag 1033
Db 392 CACAAGTGTGTTACTGTGGCGGAGCTACAAGCAGCGTAGCTCTCTCGAGGAACACAAG 451

Qy 1034 gaacgctccgagctttcttcagaacccctgacct----- 1069
Db 452 GAGCGGTGTCAACAACCTACCTCAGTGCATGGGCTCCAGAGACAGCATCTATACAGTAGTA 511

Qy 1069 --gggggagcgtgcgaagtgtgaggcaagacacacatacaagcggagatgggaagtgaaga 1126
Db 512 AAGGAAGAAACCAACCAAGTGAAGGAGGAGAGACTTAAAGCAGATGGGATCTAAGAGA 571

Qy 1127 gctctcgtcctggacagattagcaagcaaatgtggttaagcgaagaaagctcgtacctcag 1186
Db 572 GCCTTGGTGTAGACAGACTAGTAAATATGTAGCCAAACGTAAGAGACACTATGCCACAG 631

Qy 1187 aaattcatcgtgagaagcggcactgcttcgatgccaaactacaatcccggtacatgtac 1246
Db 632 AAGTTTGTAGGTGAGAAACG-----CTTCTCAACATCTCCTTCGAG 673

Qy 1247 gagaaggagaaacagatgatgcagaccggatgatggaccagcgaatcaataaagcgaatc 1306
Db 674 GGTGCGCCGGGGAACCTATGCAGCGCTCATGTCATCGACCGAGCCATCAAGTGCCTATC 733

Qy 1307 agctatctaggggctgaagccttcgccccttagtccagatccgcctgctcccaacctct 1366
Db 734 AACTACCTGGTGGCGAGTCACTCCGACCGCTCATCCAGACTCCCAACCTCCTCTGAC 793

Qy 1367 gagaaggtcccaagtcacagcagtggtaccccatagcaacttactcggggcgatagcca 1426
Db 794 ATGGGGGTATGGGGTCCATGTACCCCTCCACAAAGCCCTTCGAGAGGGCCAGCGGCTG 853

Qy 1427 atggggggcccgagagatgaaagaaacggatccctcctgccagagaagatcttgct 1486
Db 854 TCAGCCAGGAGACAGCGCAGCGCAAAATCTTGTGTGCTCGCAAACTCCAAGTCAGCCTCC 913

Qy 1487 tctgaacagaggtctgtcccccaataacagtgccagagactcacagacacagcagcaac 1546
Db 914 AGCGAAGAGGAGCGGTCTCCCGGAGGACAGTGGCCAGGATTCACCGACACGAGAGCAAC 973

Qy 1547 cagcagga-----tcgcaacatctctaccagcaagcgaagcgaagtcggtcctc 1591

```

```

Db 974 AATGAAGAGAGGAGCGAGTGTGCGGGCCCTCAGGCCCTATCTACTGACAAACCATCAGC 1033
Qy 1592 cccagggcccgcaatggatgctctcttga-----aggaggtccctcgctctctttgaa 1645
Db 1034 TCCGGGTACGTAATGACGGTGTCTGCCCTGGTGAAGAGGAGCAGCAGCGCATGAG 1093
Qy 1646 ctctcaagcccccctcccatcttgcctgagggactccatcaaatgatacaacaaagaggg 1705
Db 1094 GCCATCGGGCCAGTATCGAGATTCCCTCGAGGGGTTCAGGTGCTGAGTGGAGAGGG 1153
Qy 1706 gaggatgagtgatgttctcgtatgacacatgacacatcctctccttagattatgtatg 1765
Db 1154 GAGAGGTGAGGGCATATCGCTGTGAACATGCGCATCTCTTCTTAGACCATGTGATG 1213
Qy 1766 ttaccatccacatggggtgccatggtttcctgtagtcccttggagttaacatgtgtggc 1825
Db 1214 TACACCATCCACATGGCTGCCATGGCTTCGAGACCCCTTCGAGTCAACCTCTGTGGC 1273
Qy 1826 tategagccacgacgctgctatgagttctctctcctcctcaccatgccagagagagca 1878
Db 1274 CATCGAGTCAGGACCGTTACGAGTTTTCGTTCACACATGACCCGAGGGAGCA 1326

RESULT 9
AF092175 2309 bp mRNA VRT 29-SEP-1998
LOCUS Danio rerio ikaros mRNA, complete cds.
ACCESSION AF092175
NID G3661582
VERSION AF092175.1 GI:3661582
KEYWORDS zebrafish.
SOURCE Danio rerio
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Rasbora; Danio.

REFERENCE 1 (bases 1 to 2309)
AUTHORS Amemiya,C. and Kawasaki,H.
TITLE Characterization of zebrafish ikaros, a gene necessary for differentiation of the immune system
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2309)
AUTHORS Amemiya,C. and Kawasaki,H.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Center for Human Genetics, Boston University School of Medicine, 715 Albany Street, Boston, MA 02118, USA

FEATURES
source Location/Qualifiers
1..2309
/organism="Danio rerio"
/db_xref="taxon:7955"
/tissue_type="spleen"
86..1699
/codon_start=1
/product="ikaros"
/protein_id="AAC61763.1"
/db_xref="PID:g3661583"
/db_xref="GI:3661583"
/translation="METEAAQMSOITGRDPSMNAEGGEDODEAMPVPEDLSASTGL
QHNRDPLKPLACNIKVEARSDENGLSCMNGEAECAEDLRLIGSGAKVNGSHAG
PDSKPAAYTAGGIRLPNGKLDKDCIGVICGNPLVMVHRKSRLEERKSVLEQOKGR
PFCNCGASFTQKGLRLHKLHSGKPKFKHLNACRRDALGHLRTHSVGKGR
KCAACGRSKRSLEEKHERCHNYLOCMGLONSIIYVKEENSQNRDMPASERAL
VLDRIANNVAKRSMPQRFVGNRLSELSFSGSGELMOPHVIDAIIINSAISYLGA
SLRPLVOTSPGSADMVSPLYNLHKSQTAEGNGVSAKDSAEHLILLKSKSVKDK
GSPSPGSDSTDESNNERSAGVSTGATGLIYLTNNHAPMGRNGGLPCVREBQR
HFPALRAAGMDLSIASSEGFKVLSGDGEELRAFCIHRVFLFDHVMYTHMGCHGR
HPEPCNLGGRSODRYEFSSHTRGHRI"
BASE COUNT 656 a 508 c 572 g 573 t
ORIGIN

```

LHSGKPKPCFSCVACRRRDALTGHLRTHSVGKPHKCNKCYGRSKYQKRSLEEHEKRC
HNYLQVSHMAAGQVSHVHPMEDCKOEPIMDNNISLVPEFPAVLEKTFANNGKR
KSTTPOKFEVGEKLMRFSPDIHFDMNLTYEKEAEMLQSHMDQAINNAITVIGAALH
PLMOHAPSTIAEVAPIVSAYSOVYHPNRIERPIISRETSDSHENMDGPTSLIRPKSR
POERASPSNSCLDSTDSHSDSOSYOGNPNLPRKQSPAYMKEDVKALDAPKAP
KGLADIKYVFNGEQEIIRAFKCHRCVFLDLHVMTIHMGCYGRDPLECNICGYS
ODRYEFSSHIVRGHTFH

misc_feature

517..834

/note="encodes N-terminal zinc finger domain"

1594..1746

misc_feature

/note="encodes C-terminal zinc finger

domain/hunchback-like zinc fingers"

534 a 466 c 445 g 402 t

BASE COUNT

ORIGIN

Query Match 16.7%; Score 331.4; DB 12; Length 1847;

Best Local Similarity 58.0%; Pred. NO. 5.2e-65;

Matches 733; Conservative 0; Mismatches 446; Indels 84; Gaps 5;

QY 701 agcagaccacacagtggaagatgaactgcagctgtgcggttctctgcattagcttc 760

Db 490 ATCCGGCTTCCGAATGGTAAATGTGACGTCTGTGGCATGTTTGCATTGGGCC 549

QY 761 aacgtcttgatgttcataagcgaagccataccgcggaacccgttccagtgtaacag 820

Db 550 AATGTGCTTATGTATATAAAGGATCACACTGTGTGAGCGGCCCTTCACCTGAAC 609

QY 821 tgcggggcatctttactcagaagaagtaacctctcctcgtctaatattaaactgcacaggg 880

Db 610 TCGGAGCTTCTTTTACCACAGAGGCAACCTTCTGAGACACATAAAGTTACACTCTG 669

QY 881 gaaaacccctttaaagtgtacacctctcgaactacgcagtcgcaaaaggagagatgcgcacg 940

Db 670 GAGAAGCCCTTCAATGTCTTTCTGTAGCTATGCTGTAGAGAAGAGGACGCTCTACA 729

QY 941 ggacaccttaggacacatctgtggagagccgtacagtgtagtgagttctgcggaagaac 1000

Db 730 GGACACTCAGGACCCATCTGTGGGTAAACCTCAAGTGTAACTACTGTGGCCGAAGC 789

QY 1001 tacaagcagaagaagctccctggaggagcacaaggaacgctgcgagctttcttcagaac 1060

Db 790 TACAAGCAGCGCAGCTCACTTGGAGGAACACAGAAGCGTGTCAACAATCTCCAGAT 849

QY 1061 cctgacctgggggaagc-----tgcaagtgtg 1088

Db 850 GTCAGCATGGAGGCTCCGGGAGGCTATGAGTCACCATGTACCGCTATGGAAGATTGT 909

QY 1089 aggcaga-----cacatcaagccgagatgggaagtagagagctctc 1132

Db 910 AAGAACAGAGGCTATCATGGACACAAATATTTCTGTGTGCTTTTGAGAGACCTGCT 969

QY 1133 gtctggacagattagcaagcaatgtggctaaagcgaagagctcgtgcctcagaattc 1192

Db 970 GTCATAGAGAAGCTCAGCGCAATATGGAAGACGCAAAAGCTCCACTCCTCAGAAGTT 1029

QY 1193 atcgggtagaagcgcactgtctc-----gatgcgaactacataccccggtacatg 1243

Db 1030 GTGGGGGAAAAGCTTATGCGATTACGCTACCCAGATATTCATTTGATATGAATTAACA 1089

QY 1244 tacyagaagagacagatgatgcagaccggagatgatgacaaagccatcaataacgcc 1303

Db 1090 TATGAGAAGAGGCTGAGCTGTAGTGCAGTCTCATATGATGACCAAGCCATCAACAATGA 1149

QY 1304 atcagctatcagggtcgaagcttccgcctctagtcagactccgctgctccacc 1363

Db 1150 ATCACTACCTTGGAGCTGAGGCCCTTCACCTCTGTATGAGCATGCACCAACCAATC 1209

QY 1364 tctgagatggtcccgatcatcagcagtggtgtaccccatagc----- 1405

Db 1210 GCTGAGGTGGCCCGAGTTATAAGCTCAGCTTATTTCTCAGGTCTATCATCCAAACAGGATA 1269

QY 1405 -----acttactcggccgatatgcceaatggggcccccaggagatggaagaaacgg 1459

Db 1270 GAAAGACCCATTAGCAGGGAACAATCTGATGTCACGAAAACAATGATGGCCCCACT 1329

QY 1460 atccctctgcagagaagatcttctctgaacagagctctgtccccaataaacagtcgc 1519

Db 1330 TCTCTCATCAGACCAAGAGTCGACCCAGGAAAGAGAGGCTCTCCCCAGCAATAGCTGC 1389

QY 1520 caggactccacagacacagcagcaaccacagagatgcaccaatctctaccagcaagc 1579

Db 1390 CTCGATTCTACTGACTAGAAAAGTAGCCATGATGACCGC---CAGTCTCTACCAAGGAAC 1446

QY 1580 cagctggctctcccccagccgcccgaatggatgcctctcttgaagagagtcctcgtct 1639

Db 1447 CTGCTTAAATCCCAAGAGGAACAAGCCAGCTTACATGAAGGAGATGTCGAAGCT 1506

QY 1640 ttgaaactctcacaagccccctcccatctgcctgagggactccatcaaaagtcatcaacaaa 1699

Db 1507 TTGGATGCTACCAAGGCCCCCAAGGCTCTCTGAAGGACATCTATAAGGTTTCAATGA 1566

QY 1700 gaaggggggagtgatgagtggtttcgtatgtgaccactgccacgctcctctctcagattat 1759

Db 1567 GAAGGAGAACACATGAAGGGCCCTTCAAGTGTGAGCACTGCCGAGTCTCTTTCTAGACCAT 1626

QY 1760 gtgagtccaccatccacatgggtgcccattgttccgtgaccccttgagtgtaacatg 1819

Db 1627 GTCATGTACACCATTCACATGGGTTGCCATGCTACCGGACCCACTGGAATGCAACATC 1686

QY 1820 tdtgctatcgaagccacagatcgctatgattctctctccatcgccagagagagagcac 1879

Db 1687 TGTGCTCAGAGAAGCCAGGACCGCTACGAAATTTTCATCACACATTTGTCGAGGGGAGCAC 1746

QY 1880 aga 1882

Db 1747 ACA 1749

RESULT 8

OMU92200

LOCUS

DEFINITION

ACCESSION

NID

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..2183

/organism="Oncorhynchus mykiss"

/strain="shasta"

/db_xref="taxon:8022"

/cell_type="thymocyte"

50..1336

/gene="Ikaros"

/note="Ik-2 isoform"

Ikaros/lyf-1; alternatively spliced form missing exon 3"

/codon_start=1

OMU92200 2183 bp mRNA VRT 02-MAY-1997
Oncorhynchus mykiss Ikaros homolog (Ikaros) mRNA, Ik-2 isoform,
complete cds.

U92200
U92200.1 GI:2062739
rainbow trout.
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
Hansen, J.D.
Isolation and characterization of Ikaros homologues in the rainbow
trout
Unpublished
2 (bases 1 to 2183)
Hansen, J.D.
Direct Submission
Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for
Immunology, Grenzacherstrasse 487, Basel CH-4005, Switzerland
Location/Qualifiers
1..2183
/organism="Oncorhynchus mykiss"
/strain="shasta"
/db_xref="taxon:8022"
/cell_type="thymocyte"
50..1336
/gene="Ikaros"
/note="Ik-2 isoform; similar to mouse and human
Ikaros/lyf-1; alternatively spliced form missing exon 3"
/codon_start=1

```
gene 1. 2496
/ gene="Ikaros"
CDS 81..1649
/ gene="Ikaros"
/ note="Ik-1 isoform; similar to mouse and human
Ikaros/LyF-1; Ikaros/LyF-1 homolog"
/ codon_start=1
/ product="Ikaros homolog"
/ protein_id="AAB97474.1"
/ db_xref="PID:g2062742"
/ db_xref="GI:2062742"
/ translation="MEMEQAQEMSPGRDPPNDVSEENDEAMPEDLSASNNLO
HNRRDREGACINIKVEARDEENGLAIDMMNGEEREECAEDLRVLDASGAVNSHQ
AGPDSKGPVSSAGGIRLPNGKLCIDICGIVICIGPNVLYHKRSHGERFQCTQGA
SFTQKGLLKHLSGEKFKCHLCNYKARRDALSGLHRTSHVGRPHKCAVCGRSY
KQSSLEELHEKCHNYLQCLQNSIYTVVEESNQREDSOMGSKRALYLDRLA
NNVAKRSTDMQFVGRFSNIFSEGGPGLMOPHVIDQAINSAIYNLGAESLRPLI
QTSPTSDMGVMGSMYPLHKPPAEGHGLSKADSAENLLAKSKASSSEKDGSPHS
QDSTDESNEEKAGVGASGLLYLTNHTISGVNGLVPLVKEQORQYEAEMRASLEI
ASEGFKVLSEGEQVRAIRCEHCRILFLDHWMTIHMCHGFRDPFECNLGHRSDR
YFSSHMIRGEHR"
```

BASE COUNT 707 a 592 c 597 g 599 t 1 others
ORIGIN

Query Match 16.9%; Score 335.6; DB 4; Length 2496;
Best Local Similarity 58.0%; Pred. No. 6e-66;
Matches 711; Conservative 0; Mismatches 449; Indels 66; Gaps 4;

Qy 701 agcagaccaacagtggaagataacgtgcagctgtgcggttatctcgtcgttc 760
Db 432 ATCCGCTTCCACCGGGAAGCTCAAGTGTGATATCTGTGGGATAGTTGGATTGGCCCC 491

Qy 761 aacgtcttgatggttcataaagcagaccatccgcgcgaacccgcttcacgtgtaatacag 820
Db 492 AATGTGCTGATGTCACAAAGCAAGGACACACTGGAGAGCGCTCCATTCCAGTCACCCAG 551

Qy 821 tgcggggcatcttttaactcagaagagtaacctctctcctcctcattataaactcacacagggg 880
Db 552 TGTGGCGCTTTCTTACCAGAAAGGCAACCTGTCTACGTCTCATCAAGCTCCACTCAGGA 611

Qy 881 gaaaaacctttaagtgtcacctcgcactacacatcagatccaaagagagatgcgctcacg 940
Db 612 GAGAAAGCCCTTCAAGTGTACCTGTGCAACATATGCTGTGCGGAGAGAGCCCTCAGC 671

Qy 941 ggacaccttagcacacattctgtggagagccgtacacagtgatgattcgtgcggaagacg 1000
Db 672 GGTCACTGCTACCCACTCTGTTGGAACCCCAACAGTGTGCTTACTGTGGGGGAGC 731

Qy 1001 tacaagcagagaagctccctggagagacacaaagaaacgctgcgagctttcttcagaac 1060
Db 732 TACAAGCAGCTAGTCTCTGAGGAACACAAAGAGCGGTGTACAACTACTCTCAGTGC 791

Qy 1061 cctgacct-----gggggacgctgcaagtgtgagagca 1093
Db 792 ATGGGCTCCAGAACAGCATCTATACACTAGTAAGGAAGAAAGCAACCAAGATGAGCAG 851

Qy 1094 agacacatcaaacgagatgggaagtggagagctctcctcctggacagattagcaagc 1153
Db 852 AGGGAAGACTTAAGCCACATGGGATCTAAGAGAGCCTTGGTGTCTAGACAGACTAGCTAAT 911

Qy 1154 aatgtggttaagcaaaaagctcagatgctcagaaattcatcgttgagaagcgcaactgc 1213
Db 912 AATGTAGCAACGTAAGAGCACTATGCCACAGAAGTTTGTAGGTGAGAAACGGCTTCTCC 971

Qy 1214 ttcatgccaactcaatcccggtcatatgtacgagaaggaagcaagatgatacagacc 1273
Db 972 AACATCTCTTCGAGGTGGGCC-----GGGGGAAGTATGACAGCCT 1013

Qy 1274 cggatgatgaccaggaagcatcaataaacgcccactcagctatctatgagggtggaagccttcgc 1333
Db 1014 CATGTGTCGACGACCGCCATCAACAGTCCCATCAACTACCTGGGTGGTGGAGTCACTCCGA 1073

```
Qy 1334 cccttagtcagactcgcgctgtctccacactctgagatgggtccagtcacagcagtg 1393  
Db 1074 CCGCTCATCAGACCTCCCAACCTCCTCTGACATGGGGGTCTATGGGCTCCATGTATCCCC 1133

Qy 1394 taccatagacacttactcggcgcatatgccaatggggggcccgagagagatggaaag 1453  
Db 1134 CTCCACAAGCCCTCGAGAGGCCACGGCTGTTCACCCAAAGGACAGCGCGGCAAAAT 1193



Qy 1454 aaacggatctctcctgccagagaagatcttgcctcttgaacgaggtctgtcccccaataac 1513  
Db 1194 CTTCTGTGTCGAAAGTCCAAAGTCAGCTCCAGCGAGAAGAGAGCGGCTCCCCCAGCCAC 1253



Qy 1514 agtgcagagactccacagacacagcagagagatgcgcaaatctctaccag 1573  
Db 1254 AGTGCCAGGATTCCACCGACAGGAGCAACAATGAAGAGAAGGAGGAGGTGTGCGGGCC 1313



Qy 1574 caaagcacctggttcctccccc-----ggcccgcaatggatgcctctt 1618  
Db 1314 TCAGGCTCTACTCTACCTGACAAACACATCAGCTCCGGGTAGCTAATGCGGTGCTGCC 1373



Qy 1619 ctga-----aggaggtccctcgtcttcttgaactcctcaagccccctccatctgctg 1672  
Db 1374 CTGTGAAGGAGGAGCAGCAGCGCAGTACGAGGCCATCGGGCCAGTATCGAGATTGCC 1433



Qy 1673 agggactccatcaagtgcacaagaagagggaggtgagtgatgtgttcgatgtgac 1732  
Db 1434 TCGGAGGGGTTCAGGTGCTGAGTGGAGGGGGGAGGAGGTGAGGCATATCGCTGTGAA 1493



Qy 1733 cactgcacgtcctctcctcctagattatgatttcaaccatccacatggggtgcatggt 1792  
Db 1494 CACTGCGGCATCTCTCTCCCTAGACCATGTATGTACACCATCCACATGGGCTGCCATGGC 1553



Qy 1793 ttccgtgaccccttgagtgtaacatgtgtggttatcgaagcagcagcagctatgagttc 1852  
Db 1554 TTCGAGACCCCTTCGAGTGCACACCTCTGTGGCATCGCAGTCAGGACCGCTTACGAGTTT 1613



Qy 1853 tctctcaatcgcagagagagagca 1878  
Db 1614 TCGTCACACATGACCCGAGGGGAGCA 1639



RESULT 7  
AF044257 LOCUS 1847 bp mRNA ROD 04-FEB-1998  
DEFINITION Mus musculus multi-zinc finger protein helios mRNA, complete cds.  
ACCESSION AF044257  
NID 92829276  
VERSION AF044257.1 GI:2829276  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1847)  
AUTHORS Hahn,K., Cobb,B.S., McCarty,A.S., Brown,K.E., Klug,C.A., Lee,R.,  
Akashi,K., Weissmann,I.L., Fisher,A.G. and Smale,S.T.  
TITLE Direct Submission  
JOURNAL Submitted (22-JAN-1998) HHMI/M&I, UCLA, 675 Circle Drive South, Box  
951662, Los Angeles, CA 90095-1662, USA  
FEATURES  
source 1..1847  
/organism="Mus musculus"  
/db_xref="taxon:10090"  
CDS 178..1758  
/note="Ikaros family member"  
/codon_start=1  
/product="multi-zinc finger protein helios"  
/protein_id="AAC00513.1"  
/db_xref="PID:g2829277"  
/db_xref="GI:2829277"  
/translation="METDAIDGYITCDNELSPGEHANMAIDLTSTPNQGASPSHM  
TSPNSVKLEMSDECDROPLEIRGHDESSLEPLIESSEVADNRKVVQDLQGE  
GTRLPNGKLKCDVCGMVGIPNLVHVRSHRTPFPFHCNCGGASFTQGNLLRHK


```

MEDLINE FEATURES source	93068267	Location/Qualifiers 1. .1550 /organism="Mus musculus" /db_xref="taxon:10090" /cell_type="T cell" /dev_stage="adult" /tissue_lib="Lambda ZAP EL4" 39. .1334 /gene="Ikaro5" 39. .1334 /gene="Ikaro5" /codon_start=1 /product="Ikaro5 DNA binding protein" /protein_id="AA66193.1" /db_xref="PID:g198287" /db_xref="GI:198287"
gene	385 a	431 c 423 g 311 t
CDS		
BASE COUNT	385 a	431 c 423 g 311 t
ORIGIN		
Query Match 17.5%; Score 348; DB-12; Length 1550; Best Local Similarity 62.0%; Pred. No. 9.5e-69; Matches 716; Conservative 0; Mismatches 375; Indels 63; Gaps 8;		
QY	794	gpcgaacgcccgttcacgtgtaacatcagtcggtggcattcttttactcagaaaagtgaaacctc 853
DB	198	GGTGAACGGCCTTTCAGTGCAACCACTGCTGGGGCTCTCTTACCAGAAAGGCAACCTC 257
QY	854	ctccgtcataataactgaacacgaggggaaaaaccttttaagtgtcacctctgcaactac 913
DB	258	CTCGGCACATCAAGTGTCTACTCGGTGAGAACCTTCAATGCCATCTTTGCAACTAT 317
QY	914	gcagcgaagagagatgcgtcagcgacacacttagcacacatctctggaagagccg 973
DB	318	GCTGCGCGGAGGAGCGCCTCACCGGCACCTGAGGACGACCTCCGTGTGTAAGCCT 377
QY	974	tacaagtgtgattctgcggaagaatacagaacagagagagctccctggaggagcacaag 1033
DB	378	CACAAATGTGGATATTGTGCCGGAGCTATAAACAGCGAAGCTCTTTAGAGGAGCATAAA 437
QY	1034	gaacgctgcgagctttcttcagaacctgacctgggggagc-----ctgcaagtgtg 1087
DB	438	GAGCGATGCCACCACTACTTTGGAAGCATGGGCTTCCGGCGGTGTGCCCACTTAAG 497
QY	1088	gaggcgaagacacatcaag-----ccgagatgggaagtggagagct 1129
DB	498	GAAGAACTAACACACAGAGATGGCAGAGACCTGTGCAAGATAGGACGAGAGAGTCC 557
QY	1130	ctcgtctgacagatagcaagaatgtggttaagcgaagaaagctcgtatgcctcgaaa 1189
DB	558	CTTGTCTCTGACAGGCTGGCAAGCAATGTGCCAAACGTAAGAGCTCTATGCTCAGAAA 617
QY	1190	ttcatcgggtgagaagcggcactcttcgatgccaactacaatcccggtctacatgacag 1249
DB	618	TTTCTTGGAGCAAGTGCCCTGTGACAGATGCCCTATGACAGTCCCACTATGAGAAGGAG 677
QY	1250	aaggagaacgagatgatgcagacccggatgatggacaaacccatcaataacgcctcagc 1309
DB	678	-----GATATGATGATCCACCGTATGACCGAGCGCCATCAACATGCCATCAAC 728
QY	1310	tatctggggctgaagccttccccccttagtcacagactccgctgctcccaactctgag 1369
DB	729	TACCTGGGGCTGAGTCCCTGCGCCCATTTGGTCAGACACACCCCGGT---AGCTCCGAG 785
QY	1370	atggtcccaagtcacagcagtggtgtaccccacatcactactcggggccgatatgccaatg 1429
DB	786	GTGGTGCCAGTCATCAGCTCCATGTACCAGCTGCACAAAGCCCCCTCAGATGGCCCCCA 845
QY	1430	ggg-----gccccgagagatggaagaacacgagatcctcctccagagaag 1477
DB	846	CGGTCAACCATTCAGCAGGACCGGTGATACTTGTCTGCTGTCTGTCACGAGCCCAAG 905
QY	1478	atcttgctcttgaaacagagtgctgtcccccaataaacagtgcccagagactccacagacc 1537
DB	906	TCTGTGTCTATCGAGGAGAGGCTTCCCGAGCAACAGTGTCCAAAGACTCCACATACA 965
QY	1538	gacagcaaccacaggg---atcgccaacatctctaccagcaagaagcgaagtgtctctccc 1594
DB	966	GAGACAACCGCGGAGGAGACAGCGCGGCTTATCTACCTAACCAACACATCAACCGG 1025
QY	1595	caggccgcaatggagtccttctgaagagagtcctcctcgtctcttttgaactctcctcaag 1654
DB	1026	CATGACGCAATGGCTGGCTCTC---AAGGAGGAGCAGCGCGCTACGAGGTGTGAGG 1082
QY	1655	ccccctcccatctgctgagggactccatcaaatgatacaaaaagagggaggtgatg 1714
DB	1083	CGGGCTCAGAGAACTCGCAGGATGCCITTCGCTGTGGTCAGCAGCAGTGCGCAGCAGCTG 1142
QY	1715	gatgtgttcgatgtgacacactgcaactgctctctctctctctctctctctctctctc 1774
DB	1143	AAGGTGTACAAGTGGCAACTGCGCGCTGCTCTCTCTGGATCAGCTCATGTATACCAT 1202
QY	1775	cacatg-----gggtgccatggttccctgacatccctttgagtgtaacatgtgtgc 1825
DB	1203	CACATGGGTGCATGGCTGCCATGGCTTTCGGGATCCCTTTGAGTGTAACTATGTGTG 1262
QY	1826	tatcgaagcgaacgacgctatgattctctctcactcgcagagagagagacagagcc 1885
DB	1263	TATCAGCCAGCAGCAGGATGACGAGTTCATCCCATATCACGCGGGGAGCATCGTTAC 1322
QY	1886	atgttgaagtggc 1899
DB	1323	CACCTGAGCTAAAC 1336
RESULT	6	
LOCUS	OMU92201	
DEFINITION	Oncohyhynchus mykiss Ikaros homolog (Ikaros) mRNA, 22-JAN-1998	
ACCESSION	U92201	
NID	92062741	
VERSION	U92201.1	GI:2062741
KEYWORDS		
SOURCE	rainbow trout.	
ORGANISM	Oncohyhynchus mykiss	
REFERENCE		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;	
TITLE	Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;	
	Salmoniformes; Salmonidae; Oncohyhynchus.	
	Hansen,J.D., Strassburger,P. and Du Pasquier,L.	
	Conservation of a master hematopoietic switch gene during	
	vertebrate evolution: isolation and characterization of Ikaros from	
	teleost and amphibian species	
	Eur. J. Immunol. 27 (11), 3049-3058 (1997)	
JOURNAL	98056818	
MEDLINE	2 (bases 1 to 2496)	
REFERENCE		
AUTHORS	Hansen,J.D.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-MAR-1997) Comparative Immunol., Basel Inst. for	
FEATURES	Immunology, grenzachstrasse 487, Basel CH-4005, Switzerland	
source	1. .2496	
	/organism="Oncohyhynchus mykiss"	
	/strain="shasta"	
	/db_xref="taxon:8022"	
	/cell_type="thymocyte"	

TITLE The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse

JOURNAL J. Immunol. 156 (2), 585-592 (1996)

MEDLINE 96132984

REMARK Genbank staff at the National Library of Medicine created this entry [NCBI gibbsg 175401] from the original journal article. This sequence comes from Fig. 1B and C.

```

FEATURES
source
1. .1788
Location/Qualifiers
map location: 7130 P11.1:

```

```
source
1: 1700
/organism="Homo sapiens"
```

```

/organism=Homo sapi
db_xref=taxon:9606"

```

```

2000      /db_xref="taxo
2001      228      1788

```

gene

```
/gene="IKARC
```

```
/note="hIk1"
```

CDS	238.	.1788
-----	------	-------

/gene="IKAROS"

```
/gene= IAKKOS
/note="lymphocyte"
```

note= Lymphocyte-TestH

protein; This sequence starts at

```
/codon_start=1
```

```
/product="hikl"
```

```
/protein_id="AAB50683.1"
```

/db_xref="PID:q1911483"

/db xref="GT:1911483"
 /db_xref="ID:91311400

```
/usr_x86_64-linux-gnu/libc.so.6: /lib/x86_64-linux-gnu/libc.so.6: version 'GLIBC_2.19' not found (required by /usr_x86_64-linux-gnu/libc.so.6)
/usr_x86_64-linux-gnu/libc.so.6: /lib/x86_64-linux-gnu/libc.so.6: version 'GLIBC_2.19' not found (required by /usr_x86_64-linux-gnu/libc.so.6)
/usr_x86_64-linux-gnu/libc.so.6: /lib/x86_64-linux-gnu/libc.so.6: version 'GLIBC_2.19' not found (required by /usr_x86_64-linux-gnu/libc.so.6)
```

```
/translation="MDADEGQDMSFSSGKESPPVSDTPDEGDEPMP IPEDLSTTSGGQ
```

QSSKSDRWASNVKVTQSDDENGRACEMGECAEDLRMLDASGEKMNGSHRDQSS

ALSGVGIRLPNGKCLKDICGIIICIGPNVLMVHKRSHTGERPFQCNQCGASFTQKGNL

LRHIKLSGEKPFKCHLCNYACRRRDALTGHLRTHSVGKPHKGCYGRSYKQRTSLEE

HKERCHNYLESMGLPGTLYPVIKEETHKHEMAEDLCKIGSERSLVLDRLASNVAKRKS

SMPOKFLGDKGLSDTPYDSATYEKENEMMKSHVMDOAINNAINYLCAESLRPLVOTPP

SMF QRT ESDKSESDFIDSRATIERENETMRSHVMDQATNNAINI LGAESURFLVQIFP
GGSEVVPVTSPLYQI.HRRSEGTBRSNHSAODSAVEYI.I.I.SKAKI.VPSEFEASPSNSC

GOSEV F V I S F M I Q L A H R K S E G I F R S N H S A Q D S A V E I L L L S K A R L V P S E R E A S F S N S C
O D S T D T E S N N E F O B S G L I V I T N H I A R B A O R V S I K E E H P A Y D I I R A S E N S O D A I R I V I V S

QDSTDTESNNEEQRSGLIYLTNHIAARRAQRVSLKEEHRAVDLLRAASENSQDALRWVS
TSCFQMKVVKCEHCBIYFI DNNMYETIUMGCCUCERRDDTECNMCCYVLCDDPYETCCGHTT

TSGEQMKVYKCEHCVRVFLDHSVYTIHMGCHGFRDPFECNMGYHSQDRYEFSSHITR
COUNTING

GEHREHMS®

COUNT	450 a	511 c	497 g	330 t
-------	-------	-------	-------	-------

INTRODUCTION

every Match 20 88: Score 411 8: DB 10: Length 1788:

Very Match
20.8%; Score 411.8; DB 10; Length 1788;
Best Local Similarity: Prod No 4 100-82.

Best Local Similarity 63.3%; Pred. No. 4.1e-83;

atches 777; Conservative 0; Mismatches 387; Indels 63; Gaps 7;

700 cagcagaccaaccagtgaggagatgaactgcgacgtgtgcgggttatcctgcattagctt 759

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

564 CATTGACTTCCTAACGGAAACTAAAGTGTGATATCTGTGGGATCATTTGCATCGGGCC 623

504 CATTCGACTTCCTAACCGGAATACATAAGTGTGATATATCTGIGGGATCATTTGGATCGGGCC 623

[illegible]

/60 caacgtcttgatgggttcataagcgaagccataccggggaacgccgttccagtgtaatca 819

TITLE Cloning and sequencing of hik-1, a cDNA encoding a human homologue of mouse Ikaros/Lyf-1

JOURNAL Immunol. Lett. 49 (1-2), 139-141 (1996)

MEDLINE 96252222

REFERENCE 2 (bases 1 to 3629)

AUTHORS Nietfeld,W.

TITLE Direct Submission

JOURNAL Submitted (10-NOV-1995) Wilfried Nietfeld, Department of Virology, University of Freiburg, Institute for Medical Microbiology and Hygiene, Hermann-Herder-Strasse 11, Freiburg 79104, Germany

FEATURES

source 1..3629

organism="Homo sapiens"

db_xref="taxon:9606"

tissue_type="bone marrow"

169..1728

/gene="hik-1"

169..1728

/gene="hik-1"

note="similar to mouse Lyf-1, encoded by GenBank Accession Number S74708; similar to mouse Ikaros DNA-binding protein, Swiss-Prot Accession Number Q03267"

/codon_start=1

/product="Ikaros/Lyf-1 homolog"

/protein_id="AAC50459.1"

/db_xref="PID:91289371"

/db_xref="GI:1289371"

/translation="MDADEGQMSQVSKESPVPVSDPDEGPMPIDLSITSSGQ QSKSDRVASNVKVEQSDDEENRACEMNCEADRLMDLDSGKMGSHRQDGS LRHLKHSGEKPFKHLNACRBDALTLRLHSHVSKPKHCYKCSRSYKQSSLEE HKERCHNLYESMGLPTLYPVIKETNHEMAEDLCKIGERSLVLRDLASNVAKRS SMPQKFLGDLGSLDTPYDSSASYEKENEMKSHYMDQAINNAILYLAESLRPLVQTP PGSEVSVIPMTOLHKLPLAEGTPRSHNSAQSDAVENLLSLKALVPSEASPSN SCQSDTDFSNNEQSRGLIYLTNHIAPHARNGLSLKEHRAVDLLKLAASNSODALR VVSTSGEOMKYVKECHRVFLDHVMYTIHMGCHFRDPPECNMGVHSQDRYEFSSH ITRGEHRHMS"

BASE COUNT 917 a 1002 c 936 g 773 t 1 others

ORIGIN

Query Match 21.6%; Score 428.2; DB 11; Length 3629;

Best Local Similarity 64.0%; Pred No. 8.6e-87;

Matches 785; Conservative 0; Mismatches 388; Indels 54; Gaps 7;

Qy 700 cagcagaccacacagtggaagatgaactgcgacgtgtgcggttatcctcgattagctt 759

Db 495 CATTCGACTTCTAACGGAAACTAAAGTGTGATATCTGTGGGATCATTTGTCATCGGGCC 554

Qy 760 caactcttgatggttatacgaagccataccgagcaccgcccgtttccagttaatca 819

Db 555 CAATGTGCTCATGGTTTCAAAAAGAGCCACACTGGAGAACGGCCCTTCCAGTGAATCA 614

Qy 820 gtgcgggcatcttttactcagaaggtaacctctcgtctatattaaactcacacggg 879

Db 615 GTCCGGGGCTCATTCACCGAAGGCAACCTGCTCCGGACATCAAGCTGCATTCGGG 674

Qy 880 ggaataac 939

Db 675 GGAGAAGCCCTTCAATGCCACCTCTGCAACTACGCCCTGCCCGGAGGAGCCCTCAC 734

Qy 940 gggacaccttaggac 999

Db 735 TGCCACCTGAGGAGGACCTCGCGTGGTAAACCTCACAAATGTGGATATTGTGCCGAAG 794

Qy 1000 ctacaagcagac 1059

ph 795 CTATAACAGCAGACCTCTTTAGAGAACATAAAGAGCCCTGCCAACACTACTTGGAAAG 854

Qy 1060 cactgacct--ggggagcgtcgaagtgtggaggaagacacacacacacacacacacacac 1108

Db 855 CATGGGCGCTCCGGGCACACTGTACCCAGCTCATTAAGAAGAAACTAATCACAGTGAAT 914

Qy 1108 -----cgagatgggaagtgaagagagctctcctctgagacattagcaag 1152

Db 915 GGCAGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGTGGACACTAGCAAG 974

Qy 1153 caatgtggctaaagcgaataagctcgcctcagaaatcctcgttgtagaagcggcactg 1212

Db 975 TAAGCTGCCAAACGTAAAGCTCTATGCCCTCAGAAATTTCTTGGGACAGGG---CCT 1031

Qy 1213 ctctgatgccaaactacaatcccggtctacatgtacgagaagagacgagatgacagac 1272

Db 1032 GTCCGACACGCCCTTACGACAGCAGCGCCAGCTACGAGAAGGAGAAGAAATGATGAAGTC 1091

Qy 1273 ccggtatgatgaccaagccatcaataacgcgcctcagcttatctagggggctgaagccttcg 1332

Db 1092 CCACGTGATGACCAAGCCCATCAACACGCCCATCAACTACCTGGGGCCGAGTCCCTCGG 1151

Qy 1333 ccccttagtcagactcgcctcgtcccaacctctgagatggtccacgtcctacacagcgtg 1392

Db 1152 CCGCTGTGTGACAGCGCCCGGG---CGGTTCCGAGGTGTCCGGTCTATCAGCCCGAT 1208

Qy 1393 gtaccatagcaacttactcggcgcatatgccaatggggggcccgccagagaga----- 1446

Db 1209 GTACCAAGCTGCACAAGCGCTCGCGAGGGGACCCCGGCTCCAACCACTCGGCCAGGA 1268

Qy 1446 -----tggaagaagaaacggatcctcctcgcagagaagatcttgcctctgaacagg 1497

Db 1269 CAGCGCGCTGAGAACCTGCTGCTCTCCAAGGCCAAGTTGGTGCCTCGGAGCGGA 1328

Qy 1498 tctgtcccccataaacagtgtcccaggatccacagacacagcagcaaccacgagg---a 1554

Db 1329 GCGTCCCGGAGCAACAGCTGTCAAGACTCCACGGACACCCAGAGACCAACAGGAGACA 1388

Qy 1555 tcgcaacatctctaccagcaaaacacgtgtctctcccccagcccgcaatggatgcc 1614

Db 1389 GCGAGCGGTCTCATCTACCTGACCAACACATCGCCCGCACGCGCACACGGCTTGTG 1448

Qy 1615 tctctgaagagaggtccctcgtcttttgaactcctcaagccccctccactcctgctgag 1674

Db 1449 ---GCTCAAGAGGAGGACCGCGCTTACGACCTGCTCGCGCGCCCTCCGAGAACTCGCA 1505

Qy 1675 ggatccatcaaatgatcaaaaaagggaggtgagtgatgtgttcgatgtgacca 1734

Db 1506 GGAGCGCTCGCGTGTGCACACACAGCGGGAGCAGATGAAGGTGTACAAAGTGGACACA 1565

Qy 1735 ctgcacgctcctctcttagattatgtgatgttcacatccacatggggtgcccattggtt 1794

Db 1566 CTGCGGGTGTCTCTCTGGATCAGTCTGTACACCATCCACATGGGCTGCCACGGCTT 1625

Qy 1795 ccgtgaccccttgatgataacatgtgtggtatcgaagccacacgacgctgagttctc 1854

Db 1626 CCGTATCCTTTTGTAGTGCAACATGTGGGTACACACAGCCAGGACCGGTACGAGTTCTC 1685

Qy 1855 cttcacatcgcagagagagacacag 1881

Db 1686 GTCGCACATAACGCGAGGGAGCACCG 1712

RESULT 3

S80876

LOCUS S80876 1788 bp mRNA PRI 27-MAR-1997

DEFINITION IKAROS=hik1 (alternatively spliced) [human, Jurkat T cell line, mRNA partial, 1788 nt].

ACCESSION S80876

NID g1911482

VERSION S80876.1 GI:1911482

KEYWORDS human Jurkat T cell line.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1788)

AUTHORS Molnar,A., Wu,P., Largespada,D.A., Vorkamp,A., Scherer,S., Copeland,N.G., Jenkins,N.A., Bruns,G. and Georgopoulos,K.

source	1. .1521	/organism="Mus musculus"	
CDS	1. .>1521	/db_xref="taxon:10090"	
		/note="Ikaro's homolog; zinc finger transcription factor"	
		/codon_start=1	
		/product="atolox"	
		/protein_id="AAB58795.1"	
		/db_xref="PID:G2150044"	
		/db_xref="GI:2150044"	
		/translation="MEDIOPTVELKSTEOPLTSPDALNDYSLPKHEIENVDSRE AFANDEDESDMSKVDKDEYSDRDENIMKPEPMGDAESEMYSYAREYSDYESIKLE RHPVDNSRPTSGKMNCDVGLSCISFNVLVHKHSHTKPERFQNCQGAFTQGNL LRHKLKTEKPEFKCHLCNYACORRDLAHLTHSVKPYKCEFCGYSYKORSLEE HKERCAFQNPDLGDAASVEARHIKAEMGSRALVLDRLASNVAKRKSSMPKFI GE KRKCFDANTPGMYTEKENEMQTRMDQAINNAISYLGAERFLVOTPPAPTSEM PVISSYPIALTRADMPGAPQEMKKRILLPEKILPSERGLSPNNQSQDSYDTSNH EDRHLHQSHVVLVPOARNGMPLLEKVPSPFELLKPPICLRDSIKVLNKEGEVNDVF RCDRHVFLDYVMTIHMCGHFRDPPFCNMGYRSHDRYEFSSHIARGERRAMLK"	
BASE COUNT	433 a	399 c	399 g
ORIGIN		290 t	
Query Match	76.7%	Score 1521;	DB 12; Length 1521;
Best Local Similarity	100.0%	Pred. No. 0;	
Matches 1521; Conservative	0;	Mismatches	0; Indels
	0;	Gaps	0;
Q/	374	atggaagatatatacaacgcgactgtggagctgaaagcacgagagagagagcctctgccaca	433
Db	1	ATGGAAGATATACAACCGACTGTGGAGCTGAAAGCAGCAGGAGGAGCGCTCTGCCACA	60
Q/	434	gagagccagacgtctgaatgactacagcttgcacaaacctcatgagatagaacagtg	493
Db	61	GAGAGCCAGACGCTGTGAATGACTACAGCTTGCCTCCAAACCTCATGAGATAGAAAACGTG	120
Q/	494	gacagtagaagacccacgaatgaagacgagatgcaggaagatctcgatgaagtg	553
Db	121	GACAGTAGAAGACCCAGCCATGAGAGAGAGATGACAGAGAAATTCATGATGAAGTG	180
Q/	554	aaagatgaatacagcagacagagatgagaacattatgaagccgagcccatggagatga	613
Db	181	AAAGATGAATACACGACAGAGATGAGAACATTATGAAGCCGAGCCCATGGAGATGCA	240
Q/	614	gaagaagtgaaatgccttacactatgcaagaatacagcagactatgaacattaaag	673
Db	241	GAGAGAGTGAATGCTTACAGCTATGCAAGAGATACAGCGACTATGAAGCATTAAG	300
Q/	674	ctggagagacacgtgccttatgacaacagcagacacacacagtggaagatgaactgcac	733
Db	301	CTGGAGAGACACGTGCCCTATGACACAGCAGACACACACAGTGGGAGATGAACCTGGAC	360
Q/	734	gtgctgggttactcgtcattagcttcaacgtttgtatggttcataagcgaagccatacc	793
Db	361	GTGCTGGGTATCCTGCTTACAGCTTCAACGCTTGTATGTTTATGAAGCGAAGCCATACC	420
Q/	794	ggcgaacccggttccagtgtaacagtgccgggcatcttttactcagaagaagtaacctc	853
Db	421	GGCGAACCCCGTTCCAGTGTATACAGTGGGGGCATCTTTTACTCAGAAAGGTAACCTC	480
Q/	854	ctccgtcatattaaactgcacaggggggaaaccccttttaagtgtcacctctgcaactac	913
Db	481	CTCCGTATATTAACTGCACACGGGGGAAACCTTTTAAGTGTACCTCTGCAACTAC	540
Q/	914	gcattgcaaaaggagatgcgctcaccgggacaccttaggacacattctgtgggaagccg	973
Db	541	GCATGCCAAAGGAGAGATGCGCTCACGGACACTTAGGACACATTCTGTGGAGAGCCG	600
Q/	974	tacaagtgtagttctgcgaaagaagctacaagcagaagaagctcccttgaggaggaacag	1033
Db	601	TACAAGTGTAGTCTGCGGAAAGAACTACAAGCAGAGAAGCTCCCTGGAGAGACAAG	660
Q/	1034	gaacgctgcccagcttttcttcagaacctgacgtgggagacgctcgaagtgtggagga	1093

Db	661	GAACGCTCCGAGCTTTTCTTCAGAACCTTGACCTGGGGACGCTGCAAGTGTGGAGCA	720
Q/	1094	agacacataaagccgagatggaaagtgaagagctctctgcttgagacagatagcaagc	1153
Db	721	AGACACATCAAGCCGAGATGGAAAGTGAGAGAGCTCTGCTCTGGACAGATTAACAAC	780
Q/	1154	aatgtgctaaagcaaaaagctcgatgctcagaattcatcgttgagagcggcactgc	1213
Db	781	AATGTGGCTAAGCGAAAAAGCTCGATGCTCAGAAATTCATCGGTGAGAAGCGGCACTGC	840
Q/	1214	ttcgatgcaactacaatcccgctacatgtacgagaagagaagagatgatgagaac	1273
Db	841	TTCGATGCCAACTACAATCCCGCTACATGTACGAGAAGAGAGAGATGATGAGAGCC	900
Q/	1274	cggatgatgacaaagccatcaatcaagccatcagctactctagggctgaaacccctccgc	1333
Db	901	CGGATGATGACCAAGCCATCAATAAGCCATCAGCTATCTAGGGGCTGAAGCCTCCGC	960
Q/	1334	cccttagtccagactcgcgctcctccacctgtgagatggtcccgatcctacgagtg	1393
Db	961	CCCTTAGTCCAGACTCCGCTCTCCACCTCTGAGATGGTCCAGTCATCAGCAGTGTG	1020
Q/	1394	taccatagcacttactcggccgatgccaatggggggcccgccagagatgaaag	1453
Db	1021	TACCCATAGCACTTACTCGGGCCGATATGCCAATGGGGGCCCCGAGAGATGGAAGA	1080
Q/	1454	aaacggtactcctcgcagagaagatcttgccttcttgaacgaggtctgtcccccataac	1513
Db	1081	AAACGSGATCTCTGCGCAGAGAAGATCTTTCCTTCTGACGAGAGTCTGTCCCCCAATAC	1140
Q/	1514	agtgcaggaactcaacagacacccagacacacacacacacacacacacacacacacac	1573
Db	1141	AGTGCCAGGAGCTCCACAGACACCCAGACACACAGGATCGCCACATCTCTACACAG	1200
Q/	1574	caagccagctggtctcccccagccgacacacacacacacacacacacacacacacac	1633
Db	1201	CRAAGCCAGCTGGTCTCCGCCAGGCCCAATGGATGGCTTCTCTGAAGGAGGTCCCT	1260
Q/	1634	cgtcttttgaactcctcaagccctcccatctcgtgagggactccatcaaaagtgc	1693
Db	1261	CGCTCTTTTGAACCTCTCAAGCCCTCCATCTGCTGAGGAGCTCCATCAAGGTATC	1320
Q/	1694	aacaaagaaggaggagtgatgagtggttccgagtgacacacacacacacacacacacac	1753
Db	1321	ACAAAGAAGGGAGGTGATGATGTTTCGATGTGACCACTGCCACGCTCTCTTCCTTA	1380
Q/	1754	gattatgtagtctcaccatccacatggggtgcccaggttccgtagtccctttagtgt	1813
Db	1381	GATTATGATGTTTACCATCCACATGGGGTGCCATGGTTTCGGTATCCCTTTGAGTGT	1440
Q/	1814	aacatgtgtggtatcgaaagccacgacgctatgagttctcctctcacatcgcagagga	1873
Db	1441	AACATGTGTGGTATCGAAGCCACGATCGCTATGAGTTCTCTCTCCTCCTCAGAGGA	1500
Q/	1874	gagcacagagccatgttgaag	1894
Db	1501	GAGCAGAGCCATGTTGAAG	1521

RESULT	2
LOCUS	HSU40462
DEFINITION	Human Ikaro's/LyF-1 homolog (hik-1) mRNA, complete cds.
ACCESSION	U40462
NID	91289370
VERSION	U40462.1
KEYWORDS	GI:1289370
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 3629)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 05:21:26 : Search time 701.47 Seconds
(without alignments)
8995.016 Million cell updates/sec

Title: US-09-019-348-1
Perfect score: 1984
Sequence: 1 cagcagcgcacccgtctcg.....gaactcaaacccctctcgag 1984

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_vi.*
- 18: em_fun.*
- 19: em_htg.*
- 20: em_humi.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_vi.*
- 34: gb_htg1.*
- 35: gb_htg2.*
- 36: gb_in1.*
- 37: gb_in2.*
- 38: em_ba1.*
- 39: em_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	1521	76.7	1521 12	AF001293 Mus muscu

2	428.2	21.6	3629	11	HSU40462
3	411.8	20.8	1788	10	S80876
4	399.6	20.1	1557	4	GGIKTRF
5	348	17.5	1550	12	MUSIKAROS
6	335.6	16.9	2496	4	OMU92201
7	331.4	16.7	1847	12	AF044257
8	290.6	14.6	2183	4	OMU92200
9	260.6	13.1	2309	4	AF092175
10	226	11.4	2688	12	AB017615
11	208.6	10.5	2301	4	OMU92198
12	187.2	9.4	684	12	S74708
13	166.2	8.4	415	4	XLU92202
14	164	8.3	2079	4	OMU92199
15	145.6	7.4	986	4	AF024439
16	111.6	5.6	148640	4	AF056116
17	91.8	4.6	3369	10	HSU411806
18	91.8	4.6	3186	10	HSOZF
19	90.8	4.6	3376	3	BT02FGENE
20	90	4.5	2686	36	SPU19831
21	89.4	4.5	2582	9	HUMZIF1
22	89.4	4.5	4208	11	AF011573
23	89.4	4.5	2582	14	G28565
24	89.2	4.5	3923	10	HSU57796
25	89.2	4.5	4848	11	HUMLD512P3
26	88.6	4.5	2342	12	MMZFPP35
27	88.6	4.5	2268	12	MUSZFPP3
28	87.2	4.4	1944	9	HUMZNF8
29	85.8	4.3	1600	4	MSSNABP
30	85.6	4.3	4194	9	AB018303
31	85.6	4.3	2432	11	HSU09413
32	85.6	4.3	3326	12	RNU41164
33	85.6	4.3	110474	35	AC007494
34	85.2	4.3	4445	4	XLXF1N
35	85	4.3	945	36	OCZF2
36	84.8	4.3	789	36	OVZF1
37	84.8	4.3	1896	36	OVZF12
38	84.6	4.3	1661	12	MMZ24763
39	84.4	4.3	2615	12	MMRR2
40	84.4	4.3	819	12	MUSFFB
41	84	4.2	1434	12	MMNZFP
42	84	4.2	3492	12	MMZFPP37P1
43	84	4.2	185032	35	AC007607
44	83	4.2	2094	12	MMZFPP29
45	82.6	4.2	5471	9	AB007886

ALIGNMENTS

RESULT 1

AF001293
LOCUS AF001293 1521 bp mRNA ROD 04-JUN-1997
DEFINITION Mus musculus transcription factor aiolos mRNA, partial cds.
ACCESSION AF001293
NID 92150043
VERSION AF001293.1 GI:2150043

KEYWORDS

house mouse.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1521)

Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T.,

Gonzales, E., Wu, P., Neben, S. and Georgopoulos, K.

Aiolos, a lymphoid restricted transcription factor that interacts

with Ikaros to regulate lymphocyte differentiation

EMBO J. 17, 2004-2013 (1997)

2 (bases 1 to 1521)

Morgan, B.

Direct Submission

Submitted (24-APR-1997) CBRC, MGH East, Bldg. 149 13th Street,

Charlestown, MA 02129, USA

Location/Qualifiers